

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 13:09:29 ; Search time 24 Seconds  
(without alignments)  
19.343 Million cell updates/sec

Title: US-09-606-910E-3  
Perfect score: 41  
Sequence: 1 PDTRPAP 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	8	4	US-09-641-833-2
2	41	100.0	9	2	US-08-288-059-13
3	41	100.0	9	2	US-08-288-059-27
4	41	100.0	9	2	US-08-666-473-110
5	41	100.0	9	3	US-08-134-198E-42
6	41	100.0	9	3	US-09-043-731-9
7	41	100.0	9	3	US-09-043-731-10
8	41	100.0	9	4	US-09-593-870A-8
9	41	100.0	9	4	US-09-593-870A-29
10	41	100.0	9	4	US-09-593-870A-69
11	41	100.0	9	4	US-09-497-232-2
12	41	100.0	9	4	US-09-497-232-19
13	41	100.0	9	4	US-09-497-232-20
14	41	100.0	10	3	US-09-339-944-9
15	41	100.0	10	4	US-09-497-232-28
16	41	100.0	10	4	US-09-497-232-29
17	41	100.0	10	4	US-09-497-232-30
18	41	100.0	10	4	US-09-497-232-31
19	41	100.0	10	4	US-09-651-265-9
20	41	100.0	13	4	US-09-914-259-77
21	41	100.0	14	2	US-08-463-230A-12
22	41	100.0	15	1	US-08-122-546-18
23	41	100.0	15	2	US-08-764-938-18
24	41	100.0	15	3	US-09-131-052-18
25	41	100.0	15	3	US-09-131-053A-18
26	41	100.0	16	3	US-09-339-944-7
27	41	100.0	16	3	US-08-737-896-1

Sequence 18, Appl  
Sequence 15, Appl  
Sequence 12, Appl  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 3, Appl  
Sequence 5, Appl  
Sequence 11, Appl  
Sequence 22, Appl  
Sequence 1, Appl  
Sequence 12, Appl  
Sequence 32, Appl  
Sequence 20, Appl  
Sequence 1, Appl  
Sequence 8, Appl

28 41 100.0 16 3 US-09-043-731-18  
29 41 100.0 16 4 US-09-497-232-15  
30 41 100.0 16 4 US-09-651-265-12  
31 41 100.0 16 5 PCT-US96-09951-1  
32 41 100.0 17 3 US-08-766-521-2  
33 41 100.0 17 3 US-08-755-413-2  
34 41 100.0 19 1 US-08-039-354-3  
35 41 100.0 19 1 US-08-099-354-5  
36 41 100.0 19 2 US-08-288-059-9  
37 41 100.0 19 2 US-08-288-059-11  
38 41 100.0 19 4 US-09-593-870A-22  
39 41 100.0 20 1 US-08-328-536-1  
40 41 100.0 20 2 US-08-288-059-1  
41 41 100.0 20 2 US-08-288-059-12  
42 41 100.0 20 2 US-08-288-059-32  
43 41 100.0 20 2 US-08-902-516-20  
44 41 100.0 20 2 US-08-833-807-1  
45 41 100.0 20 2 US-08-833-807-8

## ALIGNMENTS

## RESULT 1

US-09-641-833-2  
; Sequence 2, Application US/09641833  
; Patent No. 6716966  
; GENERAL INFORMATION:  
; APPLICANT: Madiyalakan, Ragupathy  
; TITLE OF INVENTION: Therapeutic Binding Agents Against MUC-1 Antigen and  
; TITLE OF INVENTION: Methods  
; FILE REFERENCE: 107823.127  
; CURRENT APPLICATION NUMBER: US/09/641.833  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: MUC-1 antigen  
US-09-641-833-2

Query Match 100.0%; Score 41; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7  
Db 2 PDTRPAP 8

## RESULT 2

US-08-288-059-13  
; Sequence 13, Application US/08288059  
; Patent No. 5827666  
; GENERAL INFORMATION:  
; APPLICANT: FINN, OLIVERA J.  
; APPLICANT: FONTENOT, J. D.  
; APPLICANT: MONTELAHO, RONALD C.  
; TITLE OF INVENTION: SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN  
; TITLE OF INVENTION: AND MUCIN-LIKE PEPTIDES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DARY & CUSHMAN, L.L.P.  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,059
; FILING DATE: 08-AUG-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: CHAPIN, MARLANA K.
; REGISTRATION NUMBER: 35,843
; REFERENCE/DOCKET NUMBER: 61137/205204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3711
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-288-059-13

Query Match 100.0%; Score 41; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7
DB 1 PDTRPAP 7
|||||

RESULT 3
US-08-288-059-27
; Sequence 27, Application US/08288059
; Patent No. 5827666
; GENERAL INFORMATION:
; APPLICANT: FINN, OLIVERA J.
; APPLICANT: FONTENOT, J. D.
; APPLICANT: MONTELAR, RONALD C.
; TITLE OF INVENTION: SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN
; TITLE OF INVENTION: AND MUCIN-LIKE PEPTIDES, AND USES THEREOF
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,059
; FILING DATE: 08-AUG-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: CHAPIN, MARLANA K.
; REGISTRATION NUMBER: 35,843
; REFERENCE/DOCKET NUMBER: 61137/205204
; TELEPHONE: 202-861-3711
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-288-059-27

Query Match 100.0%; Score 41; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7
DB 2 PDTRPAP 8
|||||

RESULT 4
US-08-666-473-110
; Sequence 110, Application US/08666473
; Patent No. 5843713
; GENERAL INFORMATION:
; APPLICANT: YOSHIDA, Aruto
; APPLICANT: TAKEUCHI, Makoto
; TITLE OF INVENTION: PEPTIDE SEQUENCE THAT FORMS MUCIN SUGAR
; TITLE OF INVENTION: CHAIN AND TECHNIQUE FOR MODIFYING PROTEIN TO BE LINKED
; TITLE OF INVENTION: WITH MUCIN SUGAR CHAIN
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,473
; FILING DATE: 19-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP95/02238
; FILING DATE: 01-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-22101
; FILING DATE: 09-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-269111
; FILING DATE: 01-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/837
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-666-473-110

Query Match 100.0%; Score 41; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7
DB 1 PDTRPAP 7
|||||

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 29, 2004, 15:36:58 ; Search time 2046 Seconds  
(without alignments)  
124.672 Million cell updates/sec

Title: US-09-606-910E-3

Perfect score: 41

Sequence: 1 PDTRPAP 7

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 158194

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_apool/US09606910/runat\_29122004\_131140\_4695/app\_query.fasta\_1.199  
-DB=EST -QFWT=fastap -SUFFIX=50nt.rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=50  
-USER=US09606910@cgn 1 1 3437 @runat\_29122004\_131140\_4695 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsl1:\*  
9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	82.9	31	1	AI682900 wc69g04.x
2	34	82.9	32	6	CAS58531 LBA00273
3	31	75.6	31	9	TA83008Q
4	31	75.6	39	8	AZ834606 2M0117G22
5	30	73.2	22	1	AI357851 qv73g11.x
6	30	73.2	36	4	BG122441 602353382
7	30	73.2	50	1	AU105505
8	29	70.7	24	8	AZ936903 2M0193E20
9	29	70.7	40	9	BX149818 Danio rer

c	10	29	70.7	41	8	AZ937325
	11	29	70.7	46	1	AI284041
	12	29	70.7	48	9	BX949286 Arabidops
c	13	29	70.7	50	1	AU105698
c	14	29	70.7	50	1	AU105699
c	15	28	68.3	25	1	AI521520
c	16	28	68.3	27	8	AZ835367
c	17	28	68.3	32	8	AZ764538
	18	28	68.3	33	4	BI822372
	19	28	68.3	37	1	AI689454
	20	28	68.3	40	1	AA934396
	21	28	68.3	40	8	CC060546
c	22	28	68.3	44	8	AZ498888
	23	28	68.3	45	1	AV833912
c	24	28	68.3	46	1	AI098710
c	25	28	68.3	50	1	AU104946
c	26	28	68.3	50	9	CR138599
c	27	28	68.3	50	9	CC941568
c	28	28	68.3	50	9	CC941572
	29	27	65.9	23	9	TA134E10Q
c	30	27	65.9	28	1	AI358659
c	31	27	65.9	31	4	BG870909
c	32	27	65.9	33	1	AL588927
	33	27	65.9	33	4	BI684367
	34	27	65.9	34	1	AI224571
c	35	27	65.9	37	1	AA906610
	36	27	65.9	38	9	DR32H3S
	37	27	65.9	39	9	TA285C10P
	38	27	65.9	41	8	BH866295
	39	27	65.9	46	9	BX171949
c	40	27	65.9	47	8	AZ331536
	41	27	65.9	48	2	BF507210
c	42	27	65.9	49	8	BH626797
c	43	27	65.9	49	8	BH626797
	44	27	65.9	50	1	AA633761
	45	27	65.9	50	1	AU103440

ALIGNMENTS

RESULT 1

AI682900

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI682900 wc69g04.x1 NCI\_CGAP Panl Homo sapiens cDNA clone IMAGE:2323926 3'  
Similar to WP:C34D4.11 CEI7505 ;, mRNA sequence.

AI682900.1 GI:4893082  
EST.  
Homo sapiens (human)

1 (bases 1 to 31)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbe-remail.nih.gov

Life Technologies catalog #: 11548-013  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality  
Seq primer: -40UP from Gibco  
High quality sequence stop: 1.  
Location/Qualifiers  
1. .31  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

```

/clone="IMAGE:2323926"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pan1"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 3.83e+03 Length: 31
Score: 34.00 Matches: 6
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 1
Query Match: 82.93% Indels: 0
DB: 1 Gaps: 0

```

US-09-606-910E-3 (1-7) x AI682900 (1-31)

```

Qy 1 ProAspThrArgProAlaPro 7
Db 1 CCCCCCAGCGCCGCCGCCCC 21

```

## RESULT 2

```

CA585831/c
LOCUS LBA00273.BOST4 cDNA from mouse aorta Mus musculus cDNA, mRNA
DEFINITION 32 bp mRNA linear EST 12-JAN-2004
sequence.

```

```

ACCESSION CA585831
VERSION CA585831.1 GI:40790953
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

## REFERENCE

```

AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
Melville, S.E., Rajandream, M.A. and Barrell, B.G.
TITLE Vascular gene expression in atherosclerotic plaque prone regions
analysed by representational difference analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Andersson Tove
Department of Biotechnology
KTH
Teknikringen 34, plan 6, 100 44 Stockholm, Sweden
Tel: +46 8 790 71 29
Fax: +46 8 245452
Email: tove@biochem.kth.se
Genes upregulated in plaque prone regions (DP1 and DP2)
Seq primer: CTA TGA CCA TGA TTA CGC CAA G.

```

## FEATURES

## source

```

1..32
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/sex="male"
/dev_stage="8 weeks old"
/clone_lib="cDNA from mouse aorta"
/note="Organ: aorta; Site_1: DpnII; Site_2: DpnII; CDNA
was prepared from whole aorta divided in atherosclerotic
plaque prone regions (aortic arch and abdominal aorta
proximal part) and less plaque prone regions (descending
thoracic aorta and abdominal aorta distal part). CDNA was
fragmented with DpnII, linker ligated and amplified to
generated starting material for representational
difference analysis (RDA). The two cDNA pools were
subjected to iterative RDA subtraction and amplification
to enrich for gene fragments differentially expressed at
early stages of atherosclerosis."

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 3.94e+03 Length: 32
Score: 34.00 Matches: 5
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 71.43% Mismatches: 0
Query Match: 82.93% Indels: 0
DB: 6 Gaps: 0

```

US-09-606-910E-3 (1-7) x CA585831 (1-32)

```

Qy 1 ProAspThrArgProAlaPro 7
Db 23 CCTGACTCTAAGCCAGCCCT 3

```

## RESULT 3

## TA83D08Q

## LOCUS

## DEFINITION

T. brucei sheared genomic DNA clone 83d08, reverse sequence,

genomic survey sequence.

AL462389

VERSION AL462389.1 GI:11861024

KEYWORDS GSS.

SOURCE Trypanosoma brucei

ORGANISM Trypanosoma brucei

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

REFERENCE 1 (bases 1 to 31)

AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,

Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,

Melville, S.E., Rajandream, M.A. and Barrell, B.G.

Direct Submission

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing

project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,

Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and

nh@sanger.ac.uk

COMMENT Constructed at the Institute for Genomic Research (TIGR),

Rockville, MD. Genomic DNA isolated from a cloned population of

Trypanosoma brucei (TREU927/4 Gurat 10.1) was mechanically sheared

to give a tight size distribution (

4 kb). The v + i method used for the library construction is

described in detail in Smith, H. and Venter, J.C. (Making small

insert libraries for whole genome shotgun sequencing projects. In

Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org

Details of T. brucei sequencing at the Sanger Centre are available

at [http://www.sanger.ac.uk/Projects/T\\_brucei/](http://www.sanger.ac.uk/Projects/T_brucei/).

FEATURES

Location/Qualifiers

1..31

/organism="Trypanosoma brucei"

/mol\_type="genomic DNA"

/strain="TREU927"

/db\_xref="taxon:5691"

/clone="83d08"

## ORIGIN

## Alignment Scores:

```

Pred. No.: 1.25e+04 Length: 31
Score: 31.00 Matches: 5
Percent Similarity: 71.43% Conservative: 0
Best Local Similarity: 71.43% Mismatches: 2
Query Match: 75.61% Indels: 0
DB: 9 Gaps: 0

```

US-09-606-910E-3 (1-7) x TA83D08Q (1-31)

```

Qy 1 ProAspThrArgProAlaPro 7
Db 1 CCCAGTACTCGACCCACTCCC 21

```

## RESULT 4

## AZ834606/c

## LOCUS

## DEFINITION

2M0117G22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

39 bp DNA linear

GSS 20-FEB-2001

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 29, 2004, 17:05:04 ; Search time 360 Seconds  
(without alignments)  
109.827 Million cell updates/sec

Title: US-09-606-910E-3  
Perfect score: 41  
Sequence: 1 PDTRPAP 7

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4176236 seqs, 2824127955 residues

Total number of hits satisfying chosen parameters: 2810194

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US0960910/runat\_29122004\_131142\_4797/app\_query\_fasta\_1.199  
-DB=Published\_Applications\_NA -QFW=fastap -SUFFIX=50nt.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomem2  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0  
-MAXLEN=50 -USPR=US0960910@cgn\_1.1.480@runat\_29122004\_131142\_4797 -NCPU=6  
-ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	32	78.0	25	16	US-10-407-449-6
C 2	32	78.0	25	17	US-10-660-897-6
C 3	32	78.0	30	15	US-10-273-180-19
C 4	32	78.0	39	9	US-09-834-633A-37
C 5	32	78.0	47	10	US-09-993-346-74
C 6	32	78.0	50	10	US-09-993-346-95
C 7	31	75.6	25	15	US-10-098-263B-47201
C 8	31	75.6	25	15	US-10-098-263B-62272
C 9	31	75.6	25	15	US-10-098-263B-105312
C 10	31	75.6	25	15	US-10-098-263B-110205
C 11	31	75.6	25	15	US-10-098-263B-110206
C 12	31	75.6	41	16	US-10-035-833A-96
C 13	31	75.6	41	16	US-10-035-833A-251
C 14	31	75.6	41	16	US-10-035-833A-2337
C 15	31	75.6	41	16	US-10-035-833A-3742
C 16	31	75.6	41	16	US-10-035-833A-5390
C 17	31	75.6	41	16	US-10-035-833A-5390
C 18	30	73.2	17	15	US-10-238-700-3511
C 19	30	73.2	20	14	US-10-044-692-225
C 20	30	73.2	20	15	US-10-044-539-225
C 21	30	73.2	20	15	US-10-325-810-458
C 22	30	73.2	20	17	US-10-688-706-291
C 23	30	73.2	20	17	US-10-688-706-356
C 24	30	73.2	20	17	US-10-688-706-456
C 25	30	73.2	20	18	US-10-877-124-458
C 26	30	73.2	20	18	US-10-877-124-458
C 27	30	73.2	25	14	US-10-044-692-208
C 28	30	73.2	25	15	US-10-044-539-208
C 29	30	73.2	25	15	US-10-098-263B-4001
C 30	30	73.2	25	15	US-10-098-263B-5668
C 31	30	73.2	25	15	US-10-098-263B-16311
C 32	30	73.2	25	15	US-10-098-263B-43395
C 33	30	73.2	25	15	US-10-098-263B-108922
C 34	30	73.2	25	15	US-10-285-976-98
C 35	30	73.2	25	15	US-10-325-810-441
C 36	30	73.2	25	18	US-10-877-124-441
C 37	30	73.2	25	18	US-10-877-124-441
C 38	30	73.2	26	16	US-10-465-811-111
C 39	30	73.2	26	16	US-10-460-754-9
C 40	30	73.2	28	16	US-10-465-811-112
C 41	30	73.2	30	16	US-10-465-811-122
C 42	30	73.2	30	16	US-10-465-811-123
C 43	30	73.2	31	9	US-09-733-042-37
C 44	30	73.2	40	16	US-10-447-839A-78
C 45	30	73.2	40	16	US-10-447-839A-93

#### ALIGNMENTS

#### RESULT 1

US-10-407-449-6/c  
; Sequence 6, Application US/10407449  
; Publication No. US20040005601A1  
; GENERAL INFORMATION:  
; APPLICANT: Siddiqui-Jain, Adam  
; APPLICANT: Hurley, Laurence  
; APPLICANT: Farrell, Thomas  
; APPLICANT: Grand, Cory  
; APPLICANT: Bears, David  
; TITLE OF INVENTION: METHODS FOR TARGETING QUADRUPLIX DNA  
; FILE REFERENCE: 53223-20004.00  
; CURRENT APPLICATION NUMBER: US/10/407,449  
; CURRENT FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: US 60/404,966  
; PRIOR FILING DATE: 2002-08-04  
; PRIOR APPLICATION NUMBER: US 60/370,358  
; PRIOR FILING DATE: 2002-04-05  
; PRIOR APPLICATION NUMBER: Unknown  
; PRIOR FILING DATE: 2003-03-20  
; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-407-449-6

Alignment Scores:  
Pred. No.: 487 Length: 25  
Score: 32.00 Matches: 5  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 71.43% Mismatches: 1  
Query Match: 78.05% Indels: 0  
DB: 16 Gaps: 0

US-09-606-910E-3 (1-7) x US-10-407-449-6 (1-25)

QY 1 ProAspThrArgProAlaPro 7  
DB 24 CCGACTCCCGCCCTTCCT 4

RESULT 2

US-10-660-897-6/c  
; Sequence 6, Application US/10660897  
; Publication No. US20040115706A1  
; GENERAL INFORMATION:  
; APPLICANT: Jin, Cheng  
; APPLICANT: Chung, Mary  
; APPLICANT: Siddiqui-Jain, Adam  
; APPLICANT: Whitten, Jeffrey  
; APPLICANT: Farrell, Thomas

; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR IDENTIFYING  
; FILE REFERENCE: QUADRUPLUX FORMING NUCLEIC ACIDS AND MODULATORS THEREOF  
; CURRENT APPLICATION NUMBER: US/10/660,897  
; PRIOR FILING DATE: 2003-09-11  
; PRIOR APPLICATION NUMBER: 60/410,475  
; PRIOR FILING DATE: 2002-09-12

; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: quadruplex forming sequence  
US-10-660-897-6

Alignment Scores:  
Pred. No.: 487 Length: 25  
Score: 32.00 Matches: 5  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 71.43% Mismatches: 1  
Query Match: 78.05% Indels: 0  
DB: 17 Gaps: 0

US-09-606-910E-3 (1-7) x US-10-660-897-6 (1-25)

QY 1 ProAspThrArgProAlaPro 7  
DB 24 CCGACTCCCGCCCTTCCT 4

RESULT 3

US-10-273-180-19  
; Sequence 19, Application US/10273180  
; Publication No. US20030220476A1  
; GENERAL INFORMATION:  
; APPLICANT: KOH, Gou Young  
; TITLE OF INVENTION: CHIMERIC COILED COIL MOLECULES  
; FILE REFERENCE: 10010-00001  
; CURRENT APPLICATION NUMBER: US/10/273,180  
; CURRENT FILING DATE: 2002-10-18  
; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 30  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Primer for COMP  
US-10-273-180-19

Alignment Scores:  
Pred. No.: 568 Length: 30  
Score: 32.00 Matches: 5  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 71.43% Mismatches: 1  
Query Match: 78.05% Indels: 0  
DB: 15 Gaps: 0

US-09-606-910E-3 (1-7) x US-10-273-180-19 (1-30)

QY 1 ProAspThrArgProAlaPro 7  
DB 1 CCAGATCTGAGCCTAGCCCC 21

RESULT 4

US-09-894-633A-37  
; Sequence 37, Application US/09894633A  
; Patent No. US20020124285A1  
; GENERAL INFORMATION:  
; APPLICANT: Conner, Timothy  
; APPLICANT: Dubois, Patrice  
; APPLICANT: Malven, Marianne  
; APPLICANT: Masucci, James

; TITLE OF INVENTION: PLANT REGULATORY SEQUENCES FOR SELECTIVE CONTROL OF GENE EXPRESSION  
; FILE REFERENCE: 38-21 (15856)B  
; CURRENT APPLICATION NUMBER: US/09/894,633A  
; PRIOR FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: 60/214,357  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 09/894,633  
; PRIOR FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 111  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 37  
; LENGTH: 39  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(39)  
; OTHER INFORMATION: synthetic primer sequence  
US-09-894-633A-37

Alignment Scores:  
Pred. No.: 706 Length: 39  
Score: 32.00 Matches: 5  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 71.43% Mismatches: 1  
Query Match: 78.05% Indels: 0  
DB: 9 Gaps: 0

US-09-606-910E-3 (1-7) x US-09-894-633A-37 (1-39)

QY 1 ProAspThrArgProAlaPro 7  
DB 5 CCAGATCTGAGCCTAGCCCC 25

RESULT 5

US-09-993-346-74  
; Sequence 74, Application US/09993346  
; Publication No. US20030124530A1  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Cynthia A.  
; Cantor, Charles R.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 29, 2004, 15:46:38 ; Search time 69 Seconds  
(without alignments)  
72.109 Million cell updates/sec

Title: US-09-606-910E-3

Perfect score: 41

Sequence: 1 PDRPAP 7

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 905748

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/uspto\_spool/US09606910/runat 29122004 131141 4709/app query fasta\_1.199  
-DB=Issued Patents NA -QWMT=fastap -SUFFIX=50nt.rn1 -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=50  
-USER=US09606910@cgn 1.1 69 @runat 29122004 131141 4709 -NCPU=3  
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	78.0	47	1	US-08-171-389-74
2	32	78.0	47	1	US-08-123-936-74
3	32	78.0	47	2	US-08-475-228A-74
4	32	78.0	47	3	US-08-482-080A-74
5	32	78.0	47	3	US-09-354-947-74
6	32	78.0	47	5	PT-US93-12388-74
7	32	78.0	50	1	US-08-171-389-95
8	32	78.0	50	1	US-08-123-936-95
9	32	78.0	50	2	US-08-475-228A-95
10	32	78.0	50	3	US-08-482-080A-95
11	32	78.0	50	3	US-09-354-947-95
12	32	78.0	50	5	PCT-US93-12388-95

13	31	75.6	40	1	US-08-030-731A-25
c 14	31	75.6	40	1	US-08-030-731A-26
15	31	75.6	46	1	US-08-030-731A-28
c 16	31	75.6	46	1	US-08-030-731A-29
17	31	75.6	46	1	US-08-030-731A-31
c 18	31	75.6	46	1	US-08-030-731A-32
c 19	30	73.2	20	3	US-08-974-549A-458
c 20	30	73.2	20	4	US-08-912-951-225
c 21	30	73.2	20	4	US-09-402-181B-458
c 22	30	73.2	20	4	US-09-721-456-458
c 23	30	73.2	25	3	US-08-974-549A-441
c 24	30	73.2	25	4	US-08-912-951-208
c 25	30	73.2	25	4	US-08-402-181B-441
c 26	30	73.2	25	4	US-09-721-456-441
c 27	30	73.2	26	4	US-09-066-357-9
c 28	30	73.2	31	4	US-09-733-042-37
c 29	30	73.2	42	1	US-08-530-492-46
c 30	30	73.2	42	3	US-08-906-517-46
c 31	30	73.2	45	4	US-09-868-758-15
c 32	30	73.2	45	4	US-09-868-758-17
c 33	29	70.7	25	4	US-08-714-741-22
c 34	29	70.7	26	1	US-08-072-070-5
c 35	29	70.7	26	1	US-08-469-434-5
c 36	29	70.7	26	1	US-08-214-222-5
c 37	29	70.7	26	2	US-08-467-852A-6
c 38	29	70.7	26	2	US-08-468-718-5
c 39	29	70.7	26	2	US-08-247-491A-6
c 40	29	70.7	26	2	US-08-319-795-3
c 41	29	70.7	26	2	US-08-468-985-3
c 42	29	70.7	26	3	US-08-312-949-5
c 43	29	70.7	26	4	US-08-714-741-1
c 44	29	70.7	26	4	US-08-714-741-19
c 45	29	70.7	26	4	US-08-529-055-1

#### ALIGNMENTS

#### RESULT 1

US-08-171-389-74  
; Sequence 74, Application US/08171389  
; Patent No. 5578444  
; GENERAL INFORMATION:  
; APPLICANT: Edwards, Cynthia A.  
; APPLICANT: Cantor, Charles R.  
; APPLICANT: Andrews, Beth M.  
; APPLICANT: Turin, Lisa M.  
; APPLICANT: Fry, Kirk B.  
; TITLE OF INVENTION: Sequence-Directed DNA Binding  
; TITLE OF INVENTION: Molecules, Compositions and Methods  
; NUMBER OF SEQUENCES: 641  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genelabs Technologies, Inc.  
; STREET: 505 Penobscot Drive  
; CITY: Redwood City  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94063  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/171,389  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/123,936  
; FILING DATE: 17-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/996,783  
; FILING DATE: 23-DEC-1992  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/723,618  
FILING DATE: 27-JUN-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/081,070  
FILING DATE: 22-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Human complement C3 gene  
US-08-171-389-74

Alignment Scores:  
Pred. No.: 281 Length: 47  
Score: 32.00 Matches: 5  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 71.43% Mismatches: 1  
Query Match: 78.05% Indels: 0  
DB: 1 Gaps: 0

US-09-606-910E-3 (1-7) x US-08-171-389-74 (1-47)

QY 1 ProAspThrArgProAlaPro 7  
DB 17 CCAGATAAAAGCCAGCTCCA 37

RESULT 2  
US-08-123-936-74  
Sequence 74, Application US/08123936  
Patent No. 5726014  
GENERAL INFORMATION:  
APPLICANT: Edwards, Cynthia A.  
APPLICANT: Cantor, Charles R.  
APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
TITLE OF INVENTION: Screening Assay for the Detection of  
TITLE OF INVENTION: DNA-Binding Molecules  
NUMBER OF SEQUENCES: 640  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/123,936  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723,618  
FILING DATE: 27-JUN-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 47 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Human complement C3 gene  
US-08-123-936-74

Alignment Scores:  
Pred. No.: 281 Length: 47  
Score: 32.00 Matches: 5  
Percent Similarity: 85.71% Conservative: 1  
Best Local Similarity: 71.43% Mismatches: 1  
Query Match: 78.05% Indels: 0  
DB: 1 Gaps: 0

US-09-606-910E-3 (1-7) x US-08-123-936-74 (1-47)

QY 1 ProAspThrArgProAlaPro 7  
DB 17 CCAGATAAAAGCCAGCTCCA 37

RESULT 3  
US-08-475-228A-74  
Sequence 74, Application US/08475228A  
Patent No. 5869241  
GENERAL INFORMATION:  
APPLICANT: Edwards, Cynthia A.  
APPLICANT: Cantor, Charles R.  
APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
APPLICANT: Fry, Kirk E.  
TITLE OF INVENTION: Sequence-Directed DNA Binding  
TITLE OF INVENTION: Molecules, Compositions and Methods  
NUMBER OF SEQUENCES: 664  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,228A  
FILING DATE: 06-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123,936  
FILING DATE: 17-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723,618  
FILING DATE: 27-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/081,070  
FILING DATE: 22-JUN-1993  
ATTORNEY/AGENT INFORMATION:



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: December 29, 2004, 14:55:26 ; Search time 330 Seconds  
(without alignments)  
111.351 Million cell updates/sec

Title: US-09-606-910E-3  
Perfect score: 41  
Sequence: 1 PDRPAP 7

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 3992700

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2 1/USPTO.spool/US09606910/runat 29122004 131139 4673/app query.fasta\_1.199  
-DB=N Geneseq\_23Sep04 -QWTF=fastap -SUFFIX=50nt.rng -MINMATCH=0 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -ALIGN=15  
-USER=US09606910@cgn 1 1 470 @runat 29122004 131139 4673 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_23Sep04:  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	48	AA501380	AA501380 Human muc
2	34	82.9	21	AAQ04708	AAQ04708 Sequence
3	34	82.9	39	AAT90104	AAT90104 MUC1 modi
C 4	33	80.5	41	AAH46801	AAH46801 Human qui
C 5	32	78.0	25	AD86161	AD86161 BCL2 gene
6	32	78.0	30	AD123653	AD123653 Human car

ALIGNMENTS

RESULT 1

AA501380	AA501380 standard; DNA; 48 BP.
XX	
AC	AA501380;
XX	
DT	04-JUL-2001 (first entry)
XX	
DE	Human mucin peptide MUC1(16) encoding DNA sequence.
XX	
KW	Human; polymorphic epithelial cell mucin; PEM; chimeric virus particle;
KW	CVF; plant virus coat protein; comovirus; CPWV; cowpea mosaic virus;
KW	SBMV; Southern bean mosaic virus; LHSV; red clover necrotic mosaic virus;
KW	RCNMV; MUC1; tumour; cancer vaccine; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..48
FT	/*tag= a
FT	/partial
FT	/product= "MUC1(16) peptide"
FT	/note= "This sequence lacks both start and stop codons"
XX	
PN	WO200118199-A1.
PD	
XX	15-MAR-2001.
PF	
XX	11-SEP-2000; 2000WO-GB003500.
XX	

PR 09-SEP-1999; 99GB-00021337.  
XX (DOWC ) DOW CHEM CO.  
XX Bendig MM, Jones TD, Longstaff M, Hellendoorn K;  
XX WPI; 2001-244570/25.  
DR P-PSDB; AAU00488.  
XX New chimeric plant virus particles with an immunogenically active peptide  
PT of a tumor-associated mucin, useful as a vaccine or for the manufacture  
PT of a vaccine for treating and/or preventing tumors and/or cancer.  
XX Example 12; Fig 8B; 63pp; English.  
XX The present sequence encoding for human mucin peptide MUC1(16) is  
CC inserted into an immunogenic site encoding for the coat protein of  
CC several different plant viruses to create a novel chimeric virus particle  
CC (CVP). Constructs containing the MUC1(16) DNA sequence at different  
CC positions within several plant virus coat protein insertion sequences are  
CC created. MUC1(16) peptide is 1 of 7 polymorphic epithelial cell mucin  
CC (PEM) peptides (AAU00483-AAU00489) which can be used to create CVPs. The  
CC plant virus such as the comovirus CPMV (cowpea mosaic virus). The ability  
CC of the chimeric virus particle CPWV-MUC1(16) to elicit antibodies, which  
CC can cause regression of tumours expressing the MUC1 protein, is  
CC demonstrated in a mouse tumour model. Other examples of CVPs include the  
CC insertion of the mucin MUC1(16) peptide into the coat protein of other  
CC plant viruses e.g. SMV (Southern bean mosaic virus; AAU00490), LTSV  
CC (AAU00491) and RCNMV (red clover necrotic mosaic virus; AAU00492). The  
CC CVP is useful as or as part of a vaccine particularly for treating and  
CC preventing tumours and cancer. The CVP provides advantages over prior art  
CC antigen-presenting means since conventional live animal virus vectors can  
CC be avoided; as can the need for separate mucin peptide synthesis and  
CC chemical-coupling to a conventional carrier. Also, the CVP is shown to  
CC induce good mucosal immunity, and does not require the addition of  
CC exogenous adjuvants to induce a strong immune response  
XX Sequence 48 BP; 6 A; 14 C; 10 G; 18 T; 0 U; 0 Other;  
SQ

Alignment Scores: 60.9 Length: 48  
Pred. No.: 41.00 Matches: 7  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 4 Gaps: 0  
DB:

US-09-606-910E-3 (1-7) x AAS01380 (1-48)  
QY 1 ProAspThrArgProAlaPro 7  
Db 16 CCTGATACCTAGACTGCTCT 36

RESULT 2  
AAQ04708  
ID AAQ04708 standard; DNA; 21 BP.  
XX AAQ04708;  
XX  
XX 25-MAR-2003 (revised)  
DT 10-OCT-1990 (first entry)  
XX  
XX Sequence encoding epitopic fragment of polymorphic epithelial mucin  
DE (PEM).  
XX  
XX Breast cancer; cancer; polymorphic epithelial mucin; mucin; PEM.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT repeat\_region 1..3  
FT /\*tag= a

FT repeat\_region 19..21  
FT /\*tag= a  
XX  
XX WO9005142-A.  
XX  
XX 17-MAY-1990.  
XX  
XX 10-NOV-1988; 88GB-00026362.  
XX  
XX 10-NOV-1988; 88GB-00026362.  
PR 05-APR-1989; 89GB-00007860.  
XX  
XX (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.  
XX  
XX Taylorpapa J, Gendler S, Burchell J;  
XX WPI; 1990-178797/23.  
XX  
XX Polypeptide(s) having sequence of antigenic epitope on PEM - (polymorphic  
PT epithelial mucin), useful for producing antibodies for diagnosis and  
PT therapy of breast cancer.  
XX  
XX Disclosure; Page ?; -pp; English.  
XX  
XX The protein fragment encoded by the sequence lying between the two  
CC repeated regions (pdrp) is an epitope in PEM which is expressed at high  
CC levels in sufferers from breast cancer, and is thus useful in generating  
CC antibodies for diagnosis and therapy. (Updated on 25-MAR-2003 to correct  
CC PR field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-  
CC MAR-2003 to correct PI field.)  
XX  
XX Sequence 21 BP; 3 A; 11 C; 7 G; 0 T; 0 U; 0 Other;  
SQ

Alignment Scores: 366 Length: 21  
Pred. No.: 34.00 Matches: 6  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 82.93% Indels: 0  
Query Match: 2 Gaps: 0  
DB:

US-09-606-910E-3 (1-7) x AAQ04708 (1-21)  
QY 1 ProAspThrArgProAla 6  
Db 4 CCGGACACACGCGCGCC 21

RESULT 3  
AAT90104  
ID AAT90104 standard; DNA; 39 BP.  
XX  
XX AAT90104;  
XX  
XX 02-APR-1998 (first entry)  
DT  
XX  
XX MUC1 modified epitope amplifying primer PRMUCIM-5'.  
DE  
XX Hepatitis B virus core antigen; HBcAg; MUC1; tumour antigen epitope;  
XX chimeric; Ha-ras oncogene; mucin; tumour; human; PCR primer; ss.  
KW  
XX Synthetic.  
OS Homo sapiens.  
XX  
XX WO9735008-A1.  
PN  
XX 25-SEP-1997.  
PD  
XX 21-MAR-1997; 97WO-US004656.  
PF  
XX 21-MAR-1996; 96US-0013839P.  
PR  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA  
XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 29, 2004, 15:07:52 ; Search time 2542 Seconds  
(without alignments)  
130.223 Million cell updates/sec

Title: US-09-606-910E-3  
Perfect score: 41  
Sequence: 1 PDTRPAP 7

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 1785872

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09606910/runat\_29122004\_131140\_4683/app\_query.fasta.1.199  
-DB=genEmbl -QWMT=raescap -SUFFIX=50nt.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=50  
-USER=US09606910 @CGN 1.1 3731 @runat\_29122004\_131140\_4683 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.on.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	41	100.0	35	6	A21776 oligonucleo
C 2	41	100.0	35	6	AX192399 Sequence
C 3	41	100.0	35	6	BD000574 Human pol
C 4	34	82.9	21	6	A03527 Markush syn

C 5	34	82.9	30	6	AX192395 Sequence
C 6	34	82.9	30	6	BD000570 Human pol
C 7	32	78.0	37	6	CO788435 Sequence
8	32	78.0	39	6	AX356243 Sequence
9	32	78.0	47	6	AR032462 Sequence
10	32	78.0	47	6	I29202 Sequence 74
11	32	78.0	47	6	I90876 Sequence 74
12	32	78.0	47	6	AR209126 Sequence
13	32	78.0	50	6	AR032483 Sequence
14	32	78.0	50	6	I29223 Sequence 95
15	32	78.0	50	6	I90897 Sequence 95
16	32	78.0	50	6	AR209147 Sequence
17	31	75.6	40	6	I12504 Sequence 25
18	31	75.6	40	6	I12505 Sequence 26
19	31	75.6	41	6	AX513898 Sequence
20	31	75.6	41	6	AX514053 Sequence
21	31	75.6	41	6	AX516139 Sequence
22	31	75.6	41	6	AX517544 Sequence
23	31	75.6	41	6	AX519192 Sequence
24	31	75.6	41	6	AX520522 Sequence
25	31	75.6	46	6	I12506 Sequence 28
26	31	75.6	46	6	I12507 Sequence 29
27	31	75.6	46	6	I12508 Sequence 31
28	31	75.6	46	6	I12509 Sequence 32
29	30	73.2	20	6	E36911 Human telom
30	30	73.2	20	6	AR243432 Sequence
31	30	73.2	20	6	AR390588 Sequence
32	30	73.2	20	6	AR393202 Sequence
33	30	73.2	20	6	AX810493 Sequence
34	30	73.2	20	6	BD011162 Human tel
35	30	73.2	25	6	E36894 Human telom
36	30	73.2	25	6	AR243415 Sequence
37	30	73.2	25	6	AR390571 Sequence
38	30	73.2	25	6	AR393185 Sequence
39	30	73.2	25	6	AX810476 Sequence
40	30	73.2	25	6	BD011145 Human tel
41	30	73.2	26	6	CO753464 Sequence
42	30	73.2	26	6	AR374981 Sequence
43	30	73.2	27	6	BD225170 Medicinal
44	30	73.2	28	6	CO753465 Sequence
45	30	73.2	29	6	BD105752 Ciliary n

#### ALIGNMENTS

RESULT 1	A21776/c	A21776	35 bp	DNA	linear	PAT 21-JUL-1994
LOCUS	DEFINITION	oligonucleotide.				
ACCESSION	A21776					
VERSION	A21776.1	GI:583655				
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	artificial sequences.					
REFERENCE	1 (bases 1 to 35)					
AUTHORS						
TITLE	MUCIN NUCLEOTIDES					
JOURNAL	Patent: WO 9109867-A 13 11-JUL-1991;					
FEATURES	Location/Qualifiers					
source	1..35					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
ORIGIN						
Alignment Scores:						
Pred. No.:	43.5	Length:	35			
Score:	41.00	Matches:	7			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	6	Gaps:	0			

```

US-09-606-910E-3 (1-7) x A21776 (1-35)
QY 1 ProAspThrArgProAlaPro 7
DB 35 CCGGACACACAGCGCGGCCCG 15

RESULT 2
AX192399/c
LOCUS 35 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 5 from Patent EP1103623.
ACCESSION AX192399
VERSION AX192399.1 GI:15210366
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Taylor-Papadimitriou, J., Burchell, J. and Gendler, S.
AUTHORS Human mucin core protein: nucleic acid probes, peptide fragments
TITLE and antibodies thereto, and uses thereof in diagnostic and
therapeutic methods
JOURNAL Patent: EP 1103623-A 5 30-MAY-2001;
IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
FEATURES
source
Location/Qualifiers
1..35
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 43.5 Length: 35
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-606-910E-3 (1-7) x AX192399 (1-35)
QY 1 ProAspThrArgProAlaPro 7
DB 21 CCGGACACACAGCGCGGCCCG 1

RESULT 3
BD000574/c
LOCUS 35 bp DNA linear PAT 31-JAN-2002
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein.
ACCESSION BD000574
VERSION BD000574.1 GI:18623687
KEYWORDS JP 2000333675-A/5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 35)
AUTHORS Papadimitrov, J. T., Jendora, S. and Bachieru, J.
TITLE Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein
JOURNAL Patent: JP 2000333675-A 5 05-DEC-2000;
IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
COMMENT OS Homo sapiens (human)
PN JP 2000333675-A/5
PD 05-DEC-2000
PF 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269, 07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306, 09-NOV-1987 GB 8726172 PI JOYCE
TAYLOR PAPADIMITROV SANDRA JENDORA JOY BACHIERU PC
C12N15/02, A61K38/00, A61K39/395, A61K39/395, A61K49/00, A61P35/00, PC
C07K14/47,

PC C07K16/44, C12N5/10, C12P21/08//C12N15/02, C12R1:91, (C12N5/10,
C12R1:91),
PC C12N15/00, A61K37/02, C12N5/00, (C12N15/00, C12R1:91), (C12N5/00,
C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1..35
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
Location/Qualifiers
1..35
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 43.5 Length: 35
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-606-910E-3 (1-7) x BD000574 (1-35)
QY 1 ProAspThrArgProAlaPro 7
DB 21 CCGGACACACAGCGCGGCCCG 1

RESULT 4
A03527
LOCUS 21 bp DNA linear PAT 17-MAR-1993
DEFINITION Markush synthetic DNA sequence.
ACCESSION A03527
VERSION A03527.1 GI:344686
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 21)
AUTHORS POLYPEPTIDES
TITLE POLYPEPTIDES
JOURNAL Patent: WO 9005142-A 9 17-MAY-1990;
FEATURES
source
Location/Qualifiers
1..21
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

repeat_region 1..3
/region may not be present, occur once or several
times"
repeat_region 19..21
/region may not be present, occur once or several
times"

ORIGIN
Alignment Scores:
Pred. No.: 431 Length: 21
Score: 34.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 82.93% Indels: 0
DB: 6 Gaps: 0

US-09-606-910E-3 (1-7) x A03527 (1-21)
QY 1 ProAspThrArgProAla 6
DB 4 CCGGACACACAGCGGCCCG 21

RESULT 5
AX192395/c
LOCUS 30 bp DNA linear PAT 15-AUG-2001

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 13:16:09 ; Search time 89 Seconds  
(without alignments)  
45.254 Million cell updates/sec

Title: US-09-606-910E-3

Perfect score: 41

Sequence: 1 PDTRPAP 7

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 79092

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt.02.\*

1: uniprot\_sprot.\*

2: uniprot\_crembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	41	100.0	20	Q9UM18	Q9um18 homo sapien
2	29	70.7	15	MK1_PALPR	P80408 palomona pr
3	29	70.7	15	MK2A_PALPR	P80409 palomona pr
4	29	70.7	16	MK2B_PALPR	P80410 palomona pr
5	29	70.7	16	MK3_PALPR	P80411 palomona pr
6	29	70.7	33	O77492	O77492 nycticebus
7	29	70.7	43	Q6CI56	Q6ci56 yarrowia li
8	28	68.3	35	Q8FZ15	Q8fz15 brucella su
9	27	65.9	31	Q981P2	Q981p2 rhizobium l
10	27	65.9	31	P89087	P89087 hepatitis d
11	27	65.9	31	P89089	P89089 hepatitis d
12	27	65.9	43	Q14051	Q14051 homo sapien
13	27	65.9	47	Q8NDV6	Q8ndv6 homo sapien
14	27	65.9	48	TXAL_STOHE	P19651 stoichactis
15	26	63.4	20	Q8V9H3	Q8v9h3 chicken ane
16	26	63.4	26	LCLP_HUMAN	P34168 homo sapien
17	26	63.4	34	Q9QV83	Q9qvb3 rattus sp.
18	26	63.4	44	Q35081	Q35081 mus musculu
19	26	63.4	47	Q747Q0	Q747q0 geobacter s
20	26	63.4	47	AAR36606	Aar36606 geobacter
21	25	61.0	16	Q9UQ84	Q9uqs4 homo sapien
22	25	61.0	31	Q6XGD1	Q6xgdi leifsonia x
23	25	61.0	31	AAP69566	Aap69566 leifsonia
24	25	61.0	35	Q97SL6	Q97sl6 streptococc
25	25	61.0	37	Q77132	Q77132 human immun
26	25	61.0	41	Q9AD27	Q9ad27 streptomyce
27	25	61.0	42	Q16469	Q16469 homo sapien
28	25	61.0	46	Q7M1G3	Q7m1g3 lolium mult
29	25	61.0	47	P70886	P70886 butyrivibri
30	25	61.0	50	Q9TBA8	Q9tba8 mergallus a
31	25	61.0	50	Q8U5M1	Q8u5m1 agrobacteri

```

32      25      61.0      50      2      Q8YCK3      Q8yck3 brucella me
33      24.5      59.8      46      2      Q7R732      Q7r732 plasmodium
34      24      58.5      16      2      Q9TRR1      Q9trr1 oryctolagus
35      24      58.5      17      2      Q9UDD6      Q9udd6 homo sapien
36      24      58.5      19      2      Q8TB44      Q8tb44 homo sapien
37      24      58.5      22      1      WAG1_TRIWA      P24335 trimeresuru
38      24      58.5      22      1      WAG2_TRIWA      P89330 trimeresuru
39      24      58.5      22      2      Q6JD11      Q6jdl1 canis famill
40      24      58.5      23      2      Q9TRR0      Q9trr0 oryctolagus
41      24      58.5      25      2      Q69148      Q69148 human herpe
42      24      58.5      25      2      Q69474      Q69474 human herpe
43      24      58.5      28      2      Q9HDS7      Q9hdb7 homo sapien
44      24      58.5      33      2      Q9RMS9      Q9rms9 pantoea 999
45      24      58.5      34      2      Q7RXS6      Q7rxs6 neurospora

```

#### ALIGNMENTS

##### RESULT 1

```

Q9UM18      PRELIMINARY;      PRT;      20 AA.
AC      Q9UM18;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DB      Mucin (Fragment).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89235154; PubMed=2715633;
RA      King P.X., Tjandra J.J., Reynolds K., McLaughlin P.J., Purcell D.F.J.,
RA      McKenzie I.F.C.;
RT      "Reactivity of anti-human milk fat globule antibodies with synthetic
RT      peptides.";
RL      J. Immunol. 142:3503-3509(1989).
DR      EMBL; M26316; AAA36336.1; -.
DR      PIR; S10571; S10571.
FT      NON_TER      1
FT      NON_TER      20
SQ      SEQUENCE      20 AA; 1887 MW; 5B3473EABEAFAD87 CRC64;

Query Match      100.0%; Score 41; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches      7; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

Qy 1 PDTRPAP 7

Db 1 PDTRPAP 7

##### RESULT 2

```

MK1_PALPR      STANDARD;      PRT;      15 AA.
ID      MK1_PALPR
AC      P80408;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      05-JUN-2004 (Rel. 44, Last annotation update)
DB      Metanikowin I.
OS      Palomona prasina (Green shield bug).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC      Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
OC      Pentatominae; Palomona.
OX      NCBI_TaxID=55431;
RN      [1]
RP      SEQUENCE.
RT      TISSUR=Hemolymph;
RT      Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT      "The inducible antibacterial peptides of the hemipteran insect

```

```

RT Palomena prasina: identification of a unique family of proline-rich
RT peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: Antibacterial peptide active against Gram-negative
CC bacteria.
CC -!- INDUCTION: By bacterial infection.
KW Antibiotic; Direct protein sequencing; Insect immunity.
SQ SEQUENCE 15 AA; 1838 MW; 21407E663CE46299 CRC64;

Query Match 70.7%; Score 29; DB 1; Length 15;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDTRPAP 7
DB 4 PDYRPRP 10

RESULT 3
MK2A_PALPR STANDARD; PRT; 15 AA.
AC P80409;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Metanikowin IIA.
OS Palomena prasina (Green shield bug).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
OC Pentatominae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernyeh S., Cocciancich S., Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomena prasina: identification of a unique family of proline-rich
RT peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: Antibacterial peptide active against Gram-negative
CC bacteria.
CC -!- INDUCTION: By bacterial infection.
KW Antibiotic; Direct protein sequencing; Insect immunity.
SQ SEQUENCE 15 AA; 1893 MW; 23835D063B946299 CRC64;

Query Match 70.7%; Score 29; DB 1; Length 15;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDTRPAP 7
DB 4 PDYRPRP 10

RESULT 4
MK2B_PALPR STANDARD; PRT; 16 AA.
AC P80410;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Metanikowin IIB.
OS Palomena prasina (Green shield bug).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
OC Pentatominae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernyeh S., Cocciancich S., Briand J.-P., Hetru C., Bulet P.;

```

```

RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomena prasina: identification of a unique family of proline-rich
RT peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: Antibacterial peptide active against Gram-negative
CC bacteria.
CC -!- INDUCTION: By bacterial infection.
KW Antibiotic; Direct protein sequencing; Insect immunity.
SQ SEQUENCE 16 AA; 2040 MW; AF21407D063B9462 CRC64;

Query Match 70.7%; Score 29; DB 1; Length 16;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDTRPAP 7
DB 4 PDYRPRP 10

RESULT 5
MK3_PALPR STANDARD; PRT; 16 AA.
AC P80411;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Metanikowin III.
OS Palomena prasina (Green shield bug).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;
OC Pentatominae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernyeh S., Cocciancich S., Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomena prasina: identification of a unique family of proline-rich
RT peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: Antibacterial peptide active against Gram-negative
CC bacteria.
CC -!- INDUCTION: By bacterial infection.
KW Antibiotic; Direct protein sequencing; Insect immunity.
SQ SEQUENCE 16 AA; 2024 MW; A9E3835D063B9462 CRC64;

Query Match 70.7%; Score 29; DB 1; Length 16;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDTRPAP 7
DB 4 PDYRPRP 10

RESULT 6
O77492 PRELIMINARY; PRT; 33 AA.
AC O77492;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE D4 dopamine receptor (D4DR) (Fragment).
OS Nycticebus coucang (Slow loris).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Loridae; Nycticebus.
OX NCBI_TaxID=9470;
RN [1]
RP SEQUENCE FROM N.A.
RA Inoue-Murayama M., Takenaka O., Murayama Y.;
RT "Origin and divergence of tandem repeats of primate D4 dopamine
RT receptor genes.";

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 13:16:34 ; Search time 21 Seconds  
(without alignments)  
32.072 Million cell updates/sec

Title: US-09-606-910E-3  
Perfect score: 41  
Sequence: 1 PDTRPAP 7  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 11837

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	73.2	30	2 T30514	hypothetical prote
2	28	68.3	25	2 S44201	HLA-DRB1 exon2 pro
3	27	65.9	26	2 B60745	major outer membra
4	27	65.9	48	1 A30114	long neurotoxin I
5	26	63.4	26	2 S38295	serine proteinase
6	25	61.0	35	2 C95037	hypothetical prote
7	25	61.0	42	2 I70082	glycoprotein Ib al
8	25	61.0	46	2 S07073	arabinogalactan pr
9	25	61.0	47	2 T44626	hypothetical prote
10	25	61.0	49	2 B24696	tissue kallikrein
11	25	61.0	50	2 A03575	hypothetical prote
12	25	61.0	50	2 C97409	hypothetical prote
13	24	58.5	13	2 S09716	2S albumin large c
14	24	58.5	15	2 S67975	apolipoprotein Cb2
15	24	58.5	17	2 A42920	fatty acid ethyl e
16	24	58.5	19	2 B28457	proteoglycan II, b
17	24	58.5	22	2 S67974	apolipoprotein Cb1
18	24	58.5	23	2 B43836	fibronectin - rabb
19	24	58.5	24	2 B44008	lethal peptide I -
20	24	58.5	24	2 A44008	lethal peptide II
21	24	58.5	25	2 I40692	ceaa protein (IgA1
22	24	58.5	34	2 A36626	osteogenic protein
23	24	58.5	34	2 S02382	probable membrane
24	24	58.5	36	2 C82475	hypothetical prote
25	24	58.5	37	2 AG0540	hypothetical prote
26	24	58.5	41	2 C87603	hypothetical prote
27	24	58.5	41	2 A42064	lactam utilization
28	24	58.5	45	2 S24713	Ig alpha chain, tr
29	23	56.1	31	2 B18955	class II histocomp

30	23	56.1	31	2 A25629	cytochrome-c oxida
31	23	56.1	34	2 A19197	class II histocomp
32	23	56.1	36	2 C37473	type I topoisomera
33	23	56.1	36	2 T22263	hypothetical prote
34	23	56.1	46	2 T28799	hypothetical prote
35	23	56.1	47	2 JN0424	neurotoxin RTX-V -
36	23	56.1	47	2 A40703	androgen-binding p
37	23	56.1	48	1 T2AZR3	toxin RplII - sea
38	23	56.1	48	2 JN0425	neurotoxin RTX-IV
39	23	56.1	49	2 S25829	serpin III - horse
40	23	56.1	50	2 I80356	MHC HLA-DR-beta ch
41	22	53.7	19	2 S12268	Oa-2 antigen - mou
42	22	53.7	24	2 B27262	nicotinic acetylch
43	22	53.7	25	2 A48363	2-hydroxyglutaryl-
44	22	53.7	26	4 I55277	hemoglobin alpha c
45	22	53.7	30	2 S55462	mer5 protein homol

ALIGNMENTS

RESULT 1  
T30514  
hypothetical protein - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jul-2000  
C:Accession: T30514  
R:Onai, K.; Katagiri, S.; Akiyama, M.; Nakashima, H.  
Mol. Gen. Genet. 259, 264-271, 1998  
A:Title: Mutation of the gene for the second-largest subunit of RNA polymerase I prolongs  
A:Reference number: Z20838; MUID:98420098; PMID:9749669  
A:Accession: T30514  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-30 <ON>  
A:Cross-references: EMBL:AB006052; NID:G3668169; PIDN:BAA33444.1; PID:G3668170

Query Match 73.2%; Score 30; DB 2; Length 30;  
Best Local Similarity 83.3%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPA 6  
Db 11 PDTQPA 16

RESULT 2  
S44201  
HLA-DRB1 exon2 protein - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Jan-1995  
C:Accession: S44201  
R:Loeffler, D.; Kaltenebach, U.; Woelpl, A.; Eiermann, T.H.  
submitted to the EMBL Data Library, April 1994  
A:Description: Nucleotide sequence of a novel HLA-DRB1 allele, DRB1\*08??.  
A:Reference number: S44201  
A:Accession: S44201  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-25 <LO>  
A:Cross-references: EMBL:Z32685

Query Match 68.3%; Score 28; DB 2; Length 25;  
Best Local Similarity 71.4%; Pred. No. 69;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7  
Db 17 PGTAAP 23

RESULT 3  
B60745

major outer membrane protein - Fusobacterium nucleatum (ATCC 25586) (fragment)

C:Species: Fusobacterium nucleatum  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: B60745  
R:Bakken, V.; Aaro, S.; Jensen, H.B.  
J. Gen. Microbiol. 135, 3253-3262, 1989  
A:Title: Purification and partial characterization of a major outer-membrane protein of  
A:Reference number: A60745; MUID:90257576; PMID:2636259  
A:Accession: B60745  
A:Molecule type: protein  
A:Residues: 1-26 <BAK>  
C:Keywords: membrane protein

Query Match 65.9%; Score 27; DB 2; Length 26;

Best Local Similarity 71.4%; Pred. No. 1.1e+02; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
| | | | |  
DB 6 PAPRPAP 12

#### RESULT 4

A30114  
long neurotoxin I - sea anemone (Stichodactyla helianthus)  
C:Species: Stichodactyla helianthus, Stichodactis helianthus (Caribbean sea anemone)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: A30114; S47619

R:Kem, W.R.; Parten, B.; Pennington, M.W.; Price, D.A.; Dunn, B.M.

Biochemistry 28, 3483-3489, 1989  
A:Title: Isolation, characterization, and amino acid sequence of a polypeptide neurotoxin  
A:Reference number: A30114; MUID:89302949; PMID:2568126

A:Accession: A30114

A:Molecule type: protein

A:Residues: 1-48 <REM>

A:Cross-references: UNIPROT:P19651

R:Monks, S.A.; Gould, A.R.; Lumley, P.E.; Alewood, P.F.; Kem, W.R.; Goss, N.H.; Norton,

Biochim. Biophys. Acta 1207, 93-101, 1994

A:Title: Limited proteolysis study of structure-function relationships in Sh I, a polypeptide

A:Reference number: S47619; MUID:94318675; PMID:7913833

A:Accession: S47619

A:Molecule type: protein

A:Residues: 1-48 <MON>

C:Superfamily: sea anemone toxin

C:Keywords: neurotoxin

Query Match 65.9%; Score 27; DB 1; Length 48;

Best Local Similarity 71.4%; Pred. No. 2e+02; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
| | | | |  
DB 10 PDTRPAP 16

#### RESULT 5

S38295

serine proteinase - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 18-Feb-1994 #sequence\_revision 17-Nov-1995 #text\_change 09-Jul-2004

C:Accession: S38295

R:Heidtmann, H.H.; Travis, J.

Biol. Chem. Hoppe-Seyler 374, 871-875, 1993

A:Title: A novel chymotrypsin-like serine proteinase from human lung.

A:Reference number: S38295; MUID:94092341; PMID:8267879

A:Accession: S38295

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-26 <HEI>

A:Cross-references: UNIPROT:P34168

C:Superfamily: trypsin; trypsin homology

Query Match 63.4%; Score 26; DB 2; Length 26;

Best Local Similarity 80.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRP 5  
| | | | |  
DB 9 PDSRP 13

#### RESULT 6

C95037

hypothetical protein SP0316 [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C:Accession: C95037

R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heil,  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: C95037

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-35 <KUR>

A:Cross-references: UNIPROT:Q97SL6; GB:AE005672; PIDN:AAK74492.1; PID:G14971789; GSPDB:

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0316

Query Match 61.0%; Score 25; DB 2; Length 35;

Best Local Similarity 57.1%; Pred. No. 3.4e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
| | | | |  
DB 9 PNHQAPAP 15

#### RESULT 7

I70082

glycoprotein Ib alpha variant B - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 09-Jul-2004

C:Accession: I70082

R:Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.

J. Biol. Chem. 267, 10055-10061, 1992

A:Title: Polymorphism of human glycoprotein Ib alpha results from a variable number of t  
actions.

A:Reference number: I55355; MUID:92250564; PMID:1577776

A:Accession: I70082

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-42 <RES>

A:Cross-references: UNIPROT:Q16469; GB:S34439; NID:9249178; PIDN:AA822153.1; PID:G249179

C:Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein

C:Keywords: glycoprotein

Query Match 61.0%; Score 25; DB 2; Length 42;

Best Local Similarity 57.1%; Pred. No. 4.1e+02;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
| | | | |  
DB 2 PTSEAPAP 8

#### RESULT 8

S07073

arabinogalactan protein - Italian ryegrass (fragments)

C:Species: Lolium multiflorum (Italian ryegrass)

C>Date: 21-Nov-1993 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004

C:Accession: S07073



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 13:27:35 ; Search time 68 Seconds  
(without alignments)  
37.031 Million cell updates/sec

Title: US-09-606-910E-3

Perfect score: 41

Sequence: 1 PDTRPAP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 431440

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	7	14	US-10-057-136-3
2	41	100.0	8	9	US-09-994-466-2
3	41	100.0	8	15	US-10-296-317-43
4	41	100.0	12	13	US-10-101-731-1
5	41	100.0	13	14	US-10-080-608A-77
6	41	100.0	13	15	US-10-371-067-7
7	41	100.0	15	9	US-09-822-698A-40
8	41	100.0	16	9	US-09-984-183-17
9	41	100.0	16	9	US-09-984-333-7
10	41	100.0	16	11	US-09-143-379-1
11	41	100.0	18	9	US-09-815-346-3
12	41	100.0	18	14	US-10-106-876-6
13	41	100.0	18	14	US-10-106-876-17

14	41	100.0	19	9	US-09-815-346-2	Sequence 2, Appli
15	41	100.0	19	9	US-09-815-346-6	Sequence 5, Appli
16	41	100.0	19	14	US-10-106-876-5	Sequence 6, Appli
17	41	100.0	20	9	US-09-847-185-20	Sequence 20, Appli
18	41	100.0	20	9	US-09-994-466-1	Sequence 1, Appli
19	41	100.0	20	9	US-09-984-183-11	Sequence 11, Appli
20	41	100.0	20	9	US-09-984-183-16	Sequence 16, Appli
21	41	100.0	20	9	US-09-822-698A-7	Sequence 7, Appli
22	41	100.0	20	9	US-09-822-698A-8	Sequence 8, Appli
23	41	100.0	20	9	US-09-984-333-1	Sequence 1, Appli
24	41	100.0	20	9	US-09-984-333-6	Sequence 6, Appli
25	41	100.0	20	9	US-09-984-333-9	Sequence 9, Appli
26	41	100.0	20	10	US-09-834-240-1	Sequence 1, Appli
27	41	100.0	20	13	US-10-054-488-1	Sequence 1, Appli
28	41	100.0	20	14	US-10-057-136-1	Sequence 1, Appli
29	41	100.0	20	14	US-10-224-286-20	Sequence 20, Appli
30	41	100.0	20	14	US-10-335-394-40	Sequence 40, Appli
31	41	100.0	20	14	US-10-106-876-19	Sequence 19, Appli
32	41	100.0	20	14	US-10-292-896-1	Sequence 1, Appli
33	41	100.0	20	15	US-10-406-317-31	Sequence 31, Appli
34	41	100.0	20	15	US-10-612-090-3	Sequence 3, Appli
35	41	100.0	20	15	US-10-296-317-45	Sequence 45, Appli
36	41	100.0	20	15	US-10-297-168-31	Sequence 31, Appli
37	41	100.0	20	16	US-10-716-293-215	Sequence 215, App
38	41	100.0	21	14	US-10-062-710-196	Sequence 196, App
39	41	100.0	21	14	US-10-062-710-207	Sequence 207, App
40	41	100.0	21	16	US-10-380-927-1	Sequence 1, Appli
41	41	100.0	22	14	US-10-106-876-18	Sequence 18, Appli
42	41	100.0	24	14	US-10-292-896-2	Sequence 2, Appli
43	41	100.0	24	15	US-10-296-317-62	Sequence 62, Appli
44	41	100.0	25	10	US-09-881-339-3	Sequence 3, Appli
45	41	100.0	25	14	US-10-292-896-9	Sequence 9, Appli

#### ALIGNMENTS

#### RESULT 1

US-10-057-136-3  
; Sequence 3, Application US/10057136  
; Publication No. US20030021770A1  
; GENERAL INFORMATION:  
; APPLICANT: SCHLOM, JEFFREY  
; APPLICANT: KANTOR, JUDITH  
; APPLICANT: KUFE, DONALD  
; APPLICANT: PANICALI, DENNIS  
; APPLICANT: GRITZ, LINDA  
; TITLE OF INVENTION: RECOMBINANT POX VIRUS FOR IMMUNIZATION AGAINST MUC1  
; TITLE OF INVENTION: TUMOR-ASSOCIATED ANTIGEN  
; FILE REFERENCE: 700953/47113C  
; CURRENT APPLICATION NUMBER: US/10/057,136  
; CURRENT FILING DATE: 2002-01-25  
; PRIOR APPLICATION NUMBER: 09/366,670  
; PRIOR FILING DATE: 1999-08-03  
; PRIOR APPLICATION NUMBER: PCT/US98/03693  
; PRIOR FILING DATE: 1998-02-24  
; PRIOR APPLICATION NUMBER: 60/038,253  
; PRIOR FILING DATE: 1997-02-24  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-057-136-3

Query Match 100.0%; Score 41; DB 14; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7

Db 1 PDTRPAP 7

```
RESULT 2
US-09-994-466-2
; Sequence 2, Application US/09994466
; Publication No. US20020132771A1
; GENERAL INFORMATION:
; APPLICANT: Madiyalakan, R.
; TITLE OF INVENTION: THERAPEUTIC BINDING AGENTS AGAINST MUC-1 ANTIGEN AND METHODS FOR
; FILE REFERENCE: AREX-P03-002
; CURRENT APPLICATION NUMBER: US/09/994,466
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/724094
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-994-466-2

Query Match      100.0%; Score 41; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PDTRPAP 7
Db      2 PDTRPAP 8

RESULT 3
US-10-296-317-43
; Sequence 43, Application US/10296317
; Publication No. US20040057968A1
; GENERAL INFORMATION:
; APPLICANT: CEI-Sci Corp
; APPLICANT: Zimmerman, Daniel S
; APPLICANT: Sarin, Prem S
; TITLE OF INVENTION: T CELL BINDING LIGAND PEPTIDES, PEPTIDE
; FILE REFERENCE: CS-112
; CURRENT APPLICATION NUMBER: US/10/296,317
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/206548
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: PCT/US07/16793
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Muc1 Peptide M1a
US-10-296-317-43

Query Match      100.0%; Score 41; DB 15; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PDTRPAP 7
Db      2 PDTRPAP 8

RESULT 4
US-10-101-731-1
; Sequence 1, Application US/10101731
; Publication No. US20020168366A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, Michael
```

```
; APPLICANT: Person, Roland
; APPLICANT: No. US20020168366A1ajaim, Antoine
; TITLE OF INVENTION: Compositions and Methods for Producing Vascular Occlusion
; FILE REFERENCE: T57005US
; CURRENT APPLICATION NUMBER: US/10/101,731
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US/09/438,944A
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: human MUC-1
US-10-101-731-1

Query Match      100.0%; Score 41; DB 13; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PDTRPAP 7
Db      6 PDTRPAP 12

RESULT 5
US-10-080-608A-77
; Sequence 77, Application US/10080608A
; Publication No. US20030198956A1
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/10/080,608A
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Derived from Homo sapiens MUC1 epithelial mucin glycoprotein.
US-10-080-608A-77

Query Match      100.0%; Score 41; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PDTRPAP 7
Db      4 PDTRPAP 10

RESULT 6
US-10-371-067-7
; Sequence 7, Application US/10371067
; Publication No. US20040018587A1
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Williams, Mark K
; APPLICANT: Goldberg, Edward B
; TITLE OF INVENTION: Nanostructures containing antibody assembly subunits
; FILE REFERENCE: NAMP-P-003
; CURRENT APPLICATION NUMBER: US/10/371,067
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: 10/080,608
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/136,225
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 09/236,949
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 13:17:24 ; Search time 24 Seconds  
(without alignments)  
19,343 Million cell updates/sec

Title: US-09-606-910E-3

Perfect score: 41

Sequence: 1 PDTRPAP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 260538

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PTJUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	8	4	US-09-641-833-2
2	41	100.0	9	2	US-08-288-059-13
3	41	100.0	9	2	US-08-288-059-27
4	41	100.0	9	2	US-08-666-473-110
5	41	100.0	9	3	US-08-134-198E-42
6	41	100.0	9	3	US-09-043-731-9
7	41	100.0	9	3	US-09-043-731-10
8	41	100.0	9	4	US-09-593-870A-8
9	41	100.0	9	4	US-09-593-870A-29
10	41	100.0	9	4	US-09-593-870A-69
11	41	100.0	9	4	US-09-497-232-2
12	41	100.0	9	4	US-09-497-232-19
13	41	100.0	9	4	US-09-497-232-20
14	41	100.0	10	3	US-09-339-944-9
15	41	100.0	10	4	US-09-497-232-28
16	41	100.0	10	4	US-09-497-232-29
17	41	100.0	10	4	US-09-497-232-30
18	41	100.0	10	4	US-09-497-232-31
19	41	100.0	10	4	US-09-651-265-9
20	41	100.0	13	4	US-09-914-259-77
21	41	100.0	14	2	US-08-463-230A-12
22	41	100.0	15	1	US-08-122-546-18
23	41	100.0	15	3	US-08-764-938-18
24	41	100.0	15	3	US-09-131-052-18
25	41	100.0	15	3	US-09-131-053A-18
26	41	100.0	16	3	US-09-339-944-7
27	41	100.0	16	3	US-08-737-896-1

28	41	100.0	16	3	US-09-043-731-18	Sequence 18, Appl
29	41	100.0	16	4	US-09-497-232-15	Sequence 15, Appl
30	41	100.0	16	4	US-09-651-265-12	Sequence 12, Appl
31	41	100.0	17	5	PCT-US96-09951-1	Sequence 1, Appl
32	41	100.0	17	3	US-08-766-521-2	Sequence 2, Appl
33	41	100.0	17	3	US-08-755-413-2	Sequence 3, Appl
34	41	100.0	19	1	US-08-039-354-3	Sequence 5, Appl
35	41	100.0	19	1	US-08-039-354-5	Sequence 11, Appl
36	41	100.0	19	2	US-08-288-059-9	Sequence 22, Appl
37	41	100.0	19	2	US-08-288-059-11	Sequence 1, Appl
38	41	100.0	19	4	US-09-593-870A-22	Sequence 12, Appl
39	41	100.0	20	1	US-08-328-536-1	Sequence 32, Appl
40	41	100.0	20	2	US-08-288-059-1	Sequence 20, Appl
41	41	100.0	20	2	US-08-288-059-12	Sequence 1, Appl
42	41	100.0	20	2	US-08-288-059-32	Sequence 8, Appl
43	41	100.0	20	2	US-08-902-516-20	
44	41	100.0	20	2	US-08-833-807-1	
45	41	100.0	20	2	US-08-833-807-8	

#### ALIGNMENTS

##### RESULT 1

US-09-641-833-2  
; Sequence 2, Application US/09641833  
; Patent No. 6716966  
; GENERAL INFORMATION:  
; APPLICANT: Madiyalakan, Ragupathy  
; TITLE OF INVENTION: Therapeutic Binding Agents Against MUC-1 Antigen and  
; TITLE OF INVENTION: Methods  
; FILE REFERENCE: 107823.127  
; CURRENT APPLICATION NUMBER: US/09/641.833  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: MUC-1 antigen  
; US-09-641-833-2

Query Match 100.0%; Score 41; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	PDTRPAP 7
Db	2	PDTRPAP 8

##### RESULT 2

US-08-288-059-13  
; Sequence 13, Application US/08288059  
; Patent No. 5827666  
; GENERAL INFORMATION:  
; APPLICANT: FINN, OLIVERA J.  
; APPLICANT: FONTENOT, J. D.  
; APPLICANT: MONTELAPO, RONALD C.  
; TITLE OF INVENTION: SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN  
; TITLE OF INVENTION: AND MUCIN-LIKE PEPTIDES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DABBY & CUSHMAN, L.L.P.  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/288,059  
;; FILING DATE: 08-AUG-1994  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: CHAPIN, MARLANA K.  
;; REGISTRATION NUMBER: 35,843  
;; REFERENCE/DOCKET NUMBER: 61137/205204  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-861-3711  
;; TELEFAX: 202-822-0944  
;; TELEX: 6714627 CUSH  
;; INFORMATION FOR SEQ ID NO: 13:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-288-059-13

Query Match 100.0%; Score 41; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
Db 1 PDTRPAP 7

RESULT 3  
US-08-288-059-27  
; Sequence 27, Application US/08288059  
; Patent No. 5827666  
; GENERAL INFORMATION:  
; APPLICANT: FINN, OLIVERA J.  
; APPLICANT: FONTENOT, J. D.  
; TITLE OF INVENTION: MONTEJARO, RONALD C.  
; TITLE OF INVENTION: SYNTHETIC MULTIPLE TANDEM REPEAT MUCIN  
; AND MUCIN-LIKE PEPTIDES, AND USES THEREOF  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DABBY & CUSHMAN, L.L.P.  
; STREET: 1100 NEW YORK AVENUE, N.W.  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/288,059  
; FILING DATE: 08-AUG-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CHAPIN, MARLANA K.  
; REGISTRATION NUMBER: 35,843  
; REFERENCE/DOCKET NUMBER: 61137/205204  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3711  
; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid

;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-288-059-27

Query Match 100.0%; Score 41; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
Db 2 PDTRPAP 8

RESULT 4  
US-08-666-473-110  
; Sequence 110, Application US/08666473  
; Patent No. 5843713  
; GENERAL INFORMATION:  
; APPLICANT: YOSHIDA, Aruto  
; APPLICANT: TAKEUCHI, Makoto  
; TITLE OF INVENTION: PEPTIDE SEQUENCE THAT FORMS MUCIN SUGAR  
; CHAIN AND TECHNIQUE FOR MODIFYING PROTEIN TO BE LINKED  
; WITH MUCIN SUGAR CHAIN  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/666,473  
; FILING DATE: 19-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: WO PCT/JP95/02238  
; FILING DATE: 01-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-22101  
; FILING DATE: 09-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-269111  
; FILING DATE: 01-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 16887/837  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 110:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-666-473-110

Query Match 100.0%; Score 41; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
Db 1 PDTRPAP 7

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 13:13:04 ; Search time 74 Seconds  
(without alignments)  
33.934 Million cell updates/sec

Title: US-09-606-910E-3

Perfect score: 41

Sequence: 1 PDTRPAP 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 900495

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	7	2	AAY25108 Human MUC
2	41	100.0	8	2	AAY25109 Mouse IgG
3	41	100.0	8	4	AAB31286 Antigenic
4	41	100.0	8	5	AAU82061 Antigenic
5	41	100.0	8	5	AEE20933 Human tum
6	41	100.0	8	8	ADL15165 Human tum
7	41	100.0	9	2	AAR68021 Mucin pep
8	41	100.0	9	2	AAR68007 Mucin pep
9	41	100.0	9	2	AAW72723 Mucin pep
10	41	100.0	9	2	AAW72709 Mucin pep
11	41	100.0	9	2	AAY05217 MUC VNTR
12	41	100.0	9	2	AAW67597 T-cell ac
13	41	100.0	9	2	AAW67584 T-cell ac
14	41	100.0	9	2	AAW67596 T-cell ac
15	41	100.0	9	4	AAW72499 Fusion pr
16	41	100.0	9	4	AEE09568 Human muc
17	41	100.0	9	4	AEE09520 Human muc
18	41	100.0	9	4	AEE09519 Human muc
19	41	100.0	10	2	AAW67608 Human MUC
20	41	100.0	10	2	AAW67609 Human MUC
21	41	100.0	10	2	AAW67606 Human MUC
22	41	100.0	10	2	AAW67607 Human MUC
23	41	100.0	10	3	AAW80107 Anti MUC-
24	41	100.0	10	5	AEE20934 Human tum
25	41	100.0	11	4	AAB49998 Tumour an

#### ALIGNMENTS

##### RESULT 1

AAAY25108  
ID AAY25108 standard; peptide; 7 AA.

XX AC AAY25108;

XX DT 25-AUG-1999 (first entry)

XX DE Human MUC1 immunodominant epitope.

XX KW Immunodominant; epitope; MUC1; epithelial mucin; tumour; vaccine;  
mammary carcinoma; colorectal carcinoma; pancreas carcinoma.

XX OS Homo sapiens.

XX PN DE19758400-A1.

XX PD 01-JUL-1999.

XX PF 30-DEC-1997; 97DE-01058400.

XX PR 30-DEC-1997; 97DE-01058400.

XX PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.  
(HANI/) HANISCH F.

XX PI Karsten U, Hanisch F, Paulsen H;

XX DR WPI; 1999-372374/32.

XX PT New tumour vaccine - for treating tumour cells of mammary tumours,  
colorectal and pancreas carcinomas.

XX PS Example 1; Page 2; 6pp; German.

XX CC This invention describes a novel tumour vaccine which comprises a  
synthetic peptide derived from human epithelial mucin MUC1. This active  
tumour vaccine can be used against tumour cells from mammary, colorectal  
or pancreas carcinomas. This sequence represents an immunodominant  
epitope from human MUC1

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 1.7e+06;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7

Aae20935 Human tum  
Aae20936 Human tum  
Aay03870 SM3 antib  
Adc35694 Peptide e  
Ado15610 Human Bre  
Aar95561 Mucin cor  
Aay40237 Amino aci  
Aae12721 Human PH1  
Aae20939 MUC1 prot  
Aae20938 MUC1 prot  
Aae20955 MUC1 prot  
Aae20941 MUC1 prot  
Aae20942 MUC1 prot  
Aae20956 MUC1 prot  
Aae20943 MUC1 prot  
Aae20940 MUC1 prot  
Aae20957 MUC1 prot  
Abu58813 Mucin 1 (  
Aar84861 MUC-1 epi  
Aaw21978 Human MUC

```

Db      1 PDTRPAP 7
|||||
RESULT 2
AAAY25109
ID AAAY25109 standard; peptide; 8 AA.
XX
XX AC AAAY25109;
XX
XX DT 25-AUG-1999 (first entry)
XX
XX DE Mouse IgG1 epitope A76-A/C7.
XX
XX KW Immunodominant; epitope: MUC1; epithelial mucin; tumour; vaccine;
XX mammary carcinoma; colorectal carcinoma; pancreas carcinoma; murine;
XX IG1; A76-A/C7.
XX
XX OS Mus sp.
XX
XX PN DE19758400-A1.
XX
XX PD 01-JUL-1999.
XX
XX PF 30-DEC-1997; 97DE-01058400.
XX
XX PR 30-DEC-1997; 97DE-01058400.
XX
XX PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
XX (HANI/) HANISCH F.
XX
XX PI Karsten U, Hanisch F, Paulsen H;
XX
XX DR WPI; 1999-372374/32.
XX
XX PT New tumour vaccine - for treating tumour cells of mammary tumours,
XX colorectal and pancreas carcinomas.
XX
XX PS Example 1; Page 2; 6pp; German.
XX
XX CC This invention describes a novel tumour vaccine which comprises a
XX synthetic peptide derived from an immunodominant epitope from human
XX epithelial mucin MUC1. This active tumour vaccine can be used against
XX tumour cells from mammary, colorectal or pancreas carcinomas. This
XX sequence represents a murine IgG1 derived epitope, A76-A/C7
XX
XX SQ Sequence 8 AA;

Query Match 100.0%; Score 41; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7
Db 2 PDTRPAP 8

RESULT 3
AAB31286
ID AAB31286 standard; peptide; 8 AA.
XX
XX AC AAB31286;
XX
XX DT 20-APR-2001 (first entry)
XX
XX DE Antigenic peptide derived from a human MUC1 protein.
XX
XX KW MUC1; epithelial cell mucin; tumour; MUC1 ligand; cancer.
XX
XX OS Synthetic.
XX Homo sapiens.
XX
XX PN WO200077031-A2.

Query Match 100.0%; Score 41; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7
Db 2 PDTRPAP 8

RESULT 4
AAU82061
ID AAU82061 standard; peptide; 8 AA.
XX
XX AC AAU82061;
XX
XX DT 09-APR-2002 (first entry)
XX
XX DE Antigenic peptide M1a associated with cancer (Muc1).
XX
XX KW T-cell binding ligand; TCBL; peptide G'; human MHC class II beta chain;
XX peptide J; human beta-2-microglobulin; HIV-1; TCBL peptide construct;
XX immunological disorder; immune response; human immunodeficiency virus;
XX herpes simplex virus infection; HSV; malaria; tuberculosis; cancer; CEA;
XX acquired immunodeficiency syndrome; AIDS; allergy; autoimmune disease;
XX autoimmune myocarditis; cytostatic; antiinflammatory.
XX
XX OS Synthetic.
XX
XX PN WO200189286-A2.
XX
XX PD 29-NOV-2001.
XX
XX PF 24-MAY-2001; 2001WO-US016793.
XX
XX PR 24-MAY-2000; 2000US-0206548P.
XX
XX PA (CELS-) CEL-SCI CORP.
XX
XX PI Zimmerman DS, Sarin PS;
XX

```

21-DEC-2000.  
 15-JUN-2000; 2000WO-CA000711.  
 15-JUN-1999; 99US-0139263P.  
 (UYHE-) UNIV HEALTH NETWORK.  
 Garipey J, Yang S;  
 WPI; 2001-091388/10.  
 Novel ligands that binds to MUC1, an epithelial cell mucin, useful for  
 diagnosing, monitoring, treating and preventing cancer.  
 Example 2; Page 27; 60pp; English.  
 The present sequence represents a peptide derived from a human MUC1  
 protein. MUC1 is an epithelial cell mucin, which is found on, and shed  
 from, the surface of many tumours. The specification describes ligands  
 that bind to MUC1. These ligands were isolated using a phage display  
 technique using MUC1 tandem repeats as the target. The MUC1 tandem target  
 repeats comprise repeats of the peptide AAB31258. The ligands are useful  
 for detecting the presence of MUC1 in a sample, to treat or prevent  
 cancer associated with MUC1 and to prepare a medicament or diagnostic  
 agent to treat, prevent or detect cancer associated with MUC1. By  
 assaying for the binding between the MUC1 ligand and MUC1 in a sample  
 diagnosis or monitoring of cancer can be carried out. The ligands may  
 also be used to prepare antibodies  
 Sequence 8 AA;  
 Query Match 100.0%; Score 41; DB 4; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PDTRPAP 7  
 Db 2 PDTRPAP 8  
 RESULT 4  
 AAU82061  
 ID AAU82061 standard; peptide; 8 AA.  
 AC AAU82061;  
 DT 09-APR-2002 (first entry)  
 DE Antigenic peptide M1a associated with cancer (Muc1).  
 KW T-cell binding ligand; TCBL; peptide G'; human MHC class II beta chain;  
 peptide J; human beta-2-microglobulin; HIV-1; TCBL peptide construct;  
 immunological disorder; immune response; human immunodeficiency virus;  
 herpes simplex virus infection; HSV; malaria; tuberculosis; cancer; CEA;  
 acquired immunodeficiency syndrome; AIDS; allergy; autoimmune disease;  
 autoimmune myocarditis; cytostatic; antiinflammatory.  
 OS Synthetic.  
 PN WO200189286-A2.  
 PD 29-NOV-2001.  
 PF 24-MAY-2001; 2001WO-US016793.  
 PR 24-MAY-2000; 2000US-0206548P.  
 PA (CELS-) CEL-SCI CORP.  
 PI Zimmerman DS, Sarin PS;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

# ORIGIN

Alignment Scores: 2.25e+03 Length: 330  
Pred. No.: 41.00 Matches: 7  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 1

US-09-606-910E-3 (1-7) x A1925867 (1-330)

Qy 1 ProAspThrArgProAlaPro 7

Db 277 CCGGACACCGAGCGCGCCCG 257

# RESULT 4

AZ780422

LOCUS

DEFINITION AZ780422 349 bp DNA linear GSS 16-FEB-2001  
clone UUGC2M0017115 R, genomic survey sequence.

ACCESSION AZ780422 GI:12912067

VERSION GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,  
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von  
Niederhausern, A. and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: dunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0017 row: 1 column: 15  
Seq primer: CACACGAGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 349.

TITLE

JOURNAL

COMMENT

FEATURES

source

1..349  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0017115"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_libs="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel

# ORIGIN

Alignment Scores: 2.37e+03 Length: 349  
Pred. No.: 41.00 Matches: 7  
Score: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8

US-09-606-910E-3 (1-7) x AZ780422 (1-349)

Qy 1 ProAspThrArgProAlaPro 7

Db 166 CCAGACACCGAGCGCTGCTCC 186

# RESULT 5

BQ625927/c

LOCUS

DEFINITION BQ625927 362 bp mRNA linear EST 20-DEC-2002  
ph86g06.y1 Ostertagia ostertagi L3 SL1 TOPO v2 Ostertagia ostertagi  
cDNA 5', mRNA sequence.

ACCESSION BQ625927 GI:21653105

VERSION BQ625927.1

KEYWORDS EST.

SOURCE Ostertagia ostertagi

ORGANISM

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.  
1 (bases 1 to 362)  
McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,  
Wyllie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,  
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,  
Tsagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,  
Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,  
Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
McCann, R., Waterston, R. and Wilson, R.  
The Washington Univ. Nematode EST Project, 1999  
Unpublished (1999)  
Contact: McCarter JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..362  
/organism="Ostertagia ostertagi"  
/mol\_type="mRNA"  
/db\_xref="taxon:6317"  
/dev\_stage="third stage exsheathed larvae"  
/lab\_host="DH10B"  
/clone\_lib="Ostertagia ostertagi L3 SL1 TOPO v2"  
/note="Vector: pCRII-TOPO (Invitrogen); Site\_1: EcoRI;  
Site\_2: EcoRI; \*\*\*WARNING: Subsequent examination of these  
samples has revealed the presence of an additional  
Trichostrongyloidea cattle nematode, Cooperia oncophora.  
Sequences in this library may derive from either  
Ostertagia or Cooperia \*\*\* The library was constructed by  
Claire Murphy and Dr. James McCarter at Washington

electrophoresis. Vector DNA was prepared from a derivative  
of pW42 (gil14732114[gb|AF129072.1]), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

University, St. Louis. Oligo(dT)-SL1 PCR based library. Ostertagia ostertagi L3 cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. Nematodes were provided by Dr. Louis Gasbarre of the USDA, Beltsville, MD (lgasbarre@nri.barc.usda.gov). Third stage exsheathed larvae were collected from 14 day fecal-sphagnum moss cultures of Ostertagia eggs. The larvae were recovered by overnight passage on a Baermann apparatus, and then cleaned by passage through a 20 micron nylon mesh. The larvae were then subjected to a treatment with 1.25% chlorox to induce excystation. The larvae were washed with 5 changes of PBS and then pelleted and snap frozen in liquid nitrogen."

## ORIGIN

Alignment Scores:  
Pred. No.: 2.45e+03 Length: 362  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-606-910E-3 (1-7) x BQ625927 (1-362)

QY 1 ProAspThrArgProAlaPro 7  
DB 98 CCGGACACGAGCGCGCACCG 78

## RESULT 6

BQ625736/c

LOCUS BQ625736 372 bp mRNA linear EST 20-DEC-2002  
DEFINITION ph92a07.y1 Ostertagia ostertagi L3 SL1 TOPO v2 Ostertagia ostertagi cDNA 5', mRNA sequence.

ACCESSION BQ625736

VERSION BQ625736.1 GI:21652914

KEYWORDS EST

SOURCE Ostertagia ostertagi

ORGANISM Ostertagia ostertagi

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida; Trichostrongyloidea; Haemonchidae; Ostertaginae; Ostertagia.

REFERENCE 1 (bases 1 to 372)

AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,

Wyllie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

The Washington Univ. Nematode EST Project, 1999

Unpublished (1999)

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Nematodes were provided by Dr. Louis Gasbarre of the USDA, Beltsville, MD (lgasbarre@nri.barc.usda.gov).

Putative full length read

The vector to vector length is 404

Seq primer: SL1 primer.

Location/Qualifiers

1..372

/organism="Ostertagia ostertagi"

/mol\_type="mRNA"

/db\_xref="taxon:6317"

/dev\_stage="third stage exsheathed larvae"

## FEATURES

source

/lab\_host="DH10B"  
/clone\_lib="Ostertagia ostertagi L3 SL1 TOPO v2"  
/note="Vector: pCRII-TOPO (Invitrogen); Site\_1: EcoRI;  
Site\_2: EcoRI; \*\*WARNING: Subsequent examination of these samples has revealed the presence of an additional Trichostrongyloidea cattle nematode, Cooperia oncophora. Sequences in this library may derive from either Ostertagia or Cooperia.\*\* The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SL1 PCR based library. Ostertagia ostertagi L3 cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. Nematodes were provided by Dr. Louis Gasbarre of the USDA, Beltsville, MD (lgasbarre@nri.barc.usda.gov). Third stage exsheathed larvae were collected from 14 day fecal-sphagnum moss cultures of Ostertagia eggs. The larvae were recovered by overnight passage on a Baermann apparatus, and then cleaned by passage through a 20 micron nylon mesh. The larvae were then subjected to a treatment with 1.25% chlorox to induce excystation. The larvae were washed with 5 changes of PBS and then pelleted and snap frozen in liquid nitrogen."

## ORIGIN

Alignment Scores:  
Pred. No.: 2.51e+03 Length: 372  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-606-910E-3 (1-7) x BQ625736 (1-372)

QY 1 ProAspThrArgProAlaPro 7  
DB 54 CCGGACACGAGCGCGCACCA 34

## RESULT 7

BQ100139/c

LOCUS BQ100139 376 bp mRNA linear EST 10-APR-2002

DEFINITION ph73a04.y1 Ostertagia ostertagi L4 SL1 TOPO v1 Ostertagia ostertagi cDNA, mRNA sequence.

ACCESSION BQ100139

VERSION BQ100139.1 GI:20133123

KEYWORDS EST

SOURCE Ostertagia ostertagi

ORGANISM Ostertagia ostertagi

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;

Trichostrongyloidea; Haemonchidae; Ostertaginae; Ostertagia.

REFERENCE 1 (bases 1 to 376)

AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,

Wyllie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,

Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,

Tsagaris, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,

Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,

Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterston, R. and Wilson, R.

The Washington Univ. Nematode EST Project, 1999

Unpublished (1999)

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by Claire Murphy and Dr. James McCarter

at Washington University, St. Louis. Nematodes were provided by Dr.

Louis Gasbarre of the USDA, Beltsville, MD

(lgasbarre@nri.barc.usda.gov).

Putative full length read

The vector to vector length is 404

Seq primer: SL1 primer.

Location/Qualifiers

1..372

/organism="Ostertagia ostertagi"

/mol\_type="mRNA"

/db\_xref="taxon:6317"

/dev\_stage="third stage exsheathed larvae"



## University Genome Sequencing Center

Putative full length read  
The vector to vector length is 377  
Seq primer: SL1 primer.

FEATURES  
source

Location/Qualifiers  
1..376  
/organism="Ostertagia ostertagi"  
/mol\_type="mRNA"  
/db\_xref="taxon:6317"  
/dev\_stage="L4"  
/lab\_host="DH10B"  
/clone\_lib="Ostertagia ostertagi L4 SL1 TOPO v1"  
/note="Vector: pCRII-TOPO (Invitrogen); Site 1: EcoRI;  
Site 2: EcoRI; The library was constructed by Claire  
Murphy and Dr. James McCarter at Washington University,  
St. Louis. Oligo(dT)-SL1 PCR based library. Ostertagia  
ostertagi L4 cDNA PCR products of size 3400 nucleotides  
containing SL1 on the 5' end and oligo(dT) on the 3' end  
were non-directionally cloned into pCRII-TOPO (Invitrogen)  
following the Topo TA cloning protocol. Nematodes were  
provided by Dr. Louis Gasbarre of the USDA, Beltsville,  
MD (lgasbar@nri.barc.usda.gov). Late fourth stage larvae  
were recovered 10 days after infection of young calves  
with Ostertagia L3. The calves were killed and the abomasal  
(gastric glands) were recovered and washed extensively and  
then placed in abaxmann apparatus containing warm PBS for  
approximately 4 hours. The L4 that migrated from the  
tissues were washed extensively, pelleted and snap frozen  
in liquid nitrogen."

## ORIGIN

Alignment Scores:  
Pred. No.: 2.54e+03 Length: 376  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-606-910E-3 (1-7) x BQ100139 (1-376)

Qy 1 ProAspThrArgProAlaPro 7

Db 108 CCGGACACGAGCGCGCACCG 88

## RESULT 8

BU493463 378 bp mRNA linear EST 11-SEP-2002  
LOCUS Nb\_ADGR\_D3A05.T7 Nippostrongylus brasiliensis adult (full length  
enriched) Nippostrongylus brasiliensis cDNA clone Nb\_ADGR\_D3A05.5,  
similar to Y67D8B.4 CE31662 status:Confirmed TR:Q95XT9, mRNA  
sequence.

## ACCESSION

BU493463  
VERSION BU493463.1 GI:22789718  
KEYWORDS EST.

## SOURCE

ORGANISM Nippostrongylus brasiliensis  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
Trichostrongylidae; Heligmonellidae; Nippostrongylinae;  
Nippostrongylus.

## REFERENCE

1 (bases 1 to 378)  
Harcus, Y., Fernandez, C. and Maizels, R.M.  
A survey of genes expressed by Nippostrongylus brasiliensis  
Unpublished (2002)

## JOURNAL

COMMENT Contact: Maizels RM  
Institute of Cell, Animal and Population Biology  
University of Edinburgh

Ashworth Labs, West Mains Road, Edinburgh EH93JT, Scotland  
Tel: +44 131 650 5511  
Fax: +44 131 650 5450  
Email: r.maizels@ed.ac.uk

The library was prepared by Yvonne Harcus for Rick Maizels

(Edinburgh University, UK). Sequencing was also performed by  
Yvonne. The sequence contained a PolyA tail (trimmed)

## PCR Primers

FORWARD: M13F  
BACKWARD: M13R  
Plate: D3 Row: A Column: 05  
Seq primer: T7  
High quality sequence stop: 378.

## FEATURES

## source

Location/Qualifiers  
1..378  
/organism="Nippostrongylus brasiliensis"  
/mol\_type="mRNA"  
/db\_xref="taxon:27835"  
/clone="Nb\_ADGR\_D3A05"  
/sex="mixed"  
/dev\_stage="adult worm (day 6 post infection)"  
/clone\_lib="Nippostrongylus brasiliensis adult (full  
length-enriched)"  
/note="Vector: pSPORT 1; Site 1: SalI; Site 2: NotI;  
Nippostrongylus brasiliensis is a nematode parasite of  
rodents (definitive host, rat) widely used in  
immunological research. The library was constructed from  
adult worms taken from the small intestine, 6 days after  
subcutaneous inoculation of infective larvae. cDNA was  
prepared by RT-PCR and directionally cloned in a plasmid  
vector. RT was primed with a tagged oligo-dT (NotI  
primer-adaptor). This tag was then used to prime cDNA  
amplification together with a primer directed against an  
oligonucleotide previously ligated to the originally  
capped mRNA population. cDNA was ligated to SalI  
adapters, digested with NotI, and cloned into pSPORT 1  
cut with SalI and NotI."

## ORIGIN

Alignment Scores:  
Pred. No.: 2.55e+03 Length: 378  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-606-910E-3 (1-7) x BU493463 (1-378)

Qy 1 ProAspThrArgProAlaPro 7

Db 129 CCGGACACGAGACCGCGCCA 109

## RESULT 9

CL536602/c

LOCUS Bna00046 B. mallei C-5 specific library of differential DNA

DEFINITION CL536602 381 bp DNA linear GSS 20-MAY-2004  
fragments Burkholderia mallei genomic clone 1265, genomic survey  
sequence.

## ACCESSION

CL536602  
VERSION CL536602.1 GI:47531183  
KEYWORDS GSS.

## SOURCE

ORGANISM Burkholderia mallei  
Burkholderia mallei  
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
Burkholderiaceae; Burkholderia.

## REFERENCE

1 (bases 1 to 381)  
Fushan, A., Monastyrskaya, G., Abaev, I., Filyukova, O., Kostina, M.,  
Pecherskih, E. and Sverdlov, E.

## TITLE

Genome-wide identification and mapping differences between the  
genomes of B. mallei and B. pseudomallei

## JOURNAL

Unpublished (2004)

## COMMENT

Contact: Fushan AA  
Laboratory of Structure and Functions of Human Genes  
Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of  
Sciences  
16/10 Miklukho-Maklaya St., 117997 GSP, Moscow V-437, Russia  
Tel: +7 095 330 6538

Fax: +7 095 330 6538  
 Email: gal@hungen.siocb.ras.ru  
 B. mallei C-5 genomic sequence absent from B. pseudomallei C-141,  
 misc\_feature  
 Seq primer: M13 Forward  
 Class: plasmid insert  
 High quality sequence stop: 361.  
 Location/Qualifiers  
 1..381  
 /organism="Burkholderia mallei"  
 /mol\_type="genomic DNA"  
 /strain="C-5"  
 /db\_xref="taxon:13373"  
 /clone="1265"  
 /clone\_lib="B. mallei C-5 specific library of differential  
 DNA fragments"  
 /note="The library of B. mallei C-5 specific DNA fragments  
 absent from B. pseudomallei C-141 genome was obtained  
 using a version of subtractive hybridization method  
 described in [Schmidt KD, Schmidt-Rose T, Romling U,  
 Tummeler B, Differential genome analysis of bacteria by  
 genomic subtractive hybridization and pulsed field gel  
 electrophoresis. Electrophoresis. 1998 Apr;19(4):509-14.]  
 with modifications to avoid driver biotinitiation. DNAs  
 from B. mallei C-5 and B. pseudomallei C-141 were used as  
 tester and driver, respectively."

# FEATURES source

## ORIGIN

Alignment Scores:  
 Pred. No.: 2.57e+03 Length: 381  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-606-910E-3 (1-7) x CL536602 (1-381)

Qy 1 ProAspThrArgProAlaPro 7  
 |||||  
 Db 238 CCAGATACACGACCGCACCA 218

RESULT 10  
 BQ787472/c  
 LOCUS  
 DEFINITION  
 BQ787472 385 bp mRNA linear EST 26-JUL-2002  
 imlofio.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6034722  
 5' similar to TR:Q9ULD0 Q9ULD0 KIAA1290 PROTEIN ;, mRNA sequence.

ACCESSION  
 BQ787472  
 VERSION  
 BQ787472.1 GI:21995944  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 385)  
 REFERENCE  
 AUTHORS  
 Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,  
 Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,  
 Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blisstein,A.,  
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,  
 Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,  
 Williams,T., Jackson,Y. and Bowers,Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)

TITLE  
 JOURNAL  
 COMMENT  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138

Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
 pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington

University Genome Sequencing Center For information on obtaining a  
 clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 378.

# FEATURES source

Location/Qualifiers  
 1..385  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6034722"  
 /tissue\_type="insulinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Human insulinoma"  
 /note="Organ: pancreas; Vector: pBluescript SK-; Site 1:  
 XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system  
 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
 pBluescript SK- by Dr. H. Inoue following the Washington  
 University protocol  
 (http://genome.wustl.edu/est/lambda\_protocol.shtml).  
 please contact Hiroshi Inoue, MD/PhD for further  
 information on this library (Metabolism Division, Permutt  
 Laboratory, Washington University School of Medicine, Box  
 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
 is a Washington University Pancreas EST project library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2.6e+03 Length: 385  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-09-606-910E-3 (1-7) x BQ787472 (1-385)

Qy 1 ProAspThrArgProAlaPro 7  
 |||||  
 Db 118 CCTGACACCGAGCCAGCTCCG 98

RESULT 11  
 BQ626165/c  
 LOCUS

DEFINITION  
 BQ626165 396 bp mRNA linear EST 20-DEC-2002  
 ph94c06.y1 Ostertagia ostertagi L3 SL1 TOPO v2 Ostertagia ostertagi  
 CDNA 5', mRNA sequence.

ACCESSION  
 BQ626165  
 VERSION  
 BQ626165.1 GI:21653343  
 KEYWORDS  
 EST.  
 SOURCE  
 Ostertagia ostertagi

ORGANISM  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.

REFERENCE  
 AUTHORS

McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,  
 Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,F., Theising,B.,  
 Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,  
 Tsagarishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,  
 Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,  
 Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
 McCann,R., Waterston,R. and Wilson,R.  
 The Washington Univ. Nematode EST Project, 1999  
 Unpublished (1999)

TITLE  
 JOURNAL  
 COMMENT

Contact: McCarter JP  
 The Washington Univ. Nematode EST Project, 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by Claire Murphy and Dr. James McCarter  
 at Washington University, St. Louis. Nematodes were provided by Dr.  
 Louis Gasbarrre of the USDA, Beltsville, MD  
 (lgasbarr@nri.barc.usda.gov).

Seq primer: SL1 primer  
High quality sequence stop: 391.  
Location/Qualifiers  
1. .396  
FEATURES  
source  
/organism="Ostertagia ostertagi"  
/mol\_type="mRNA"  
/db\_xref="taxon:6317"  
/dev\_stage="third stage exsheathed larvae"  
/lab\_host="DH10B"  
/clone\_lib="Ostertagia ostertagi L3 SL1 TOPO v2"  
/note="Vector: pCRII-TOPO (Invitrogen); Site 1: EcoRI;  
Site 2: EcoRI; \*\*\*WARNING: Subsequent examination of these  
samples has revealed the presence of an additional  
Trichostrongyloidea cattle nematode, Cooperia oncophora.  
Sequences in this library may derive from either  
Ostertagia or Cooperia.\*\*\* The library was constructed by  
Claire Murphy and Dr. James McCarter at Washington  
University, St. Louis. Oligo(dT)-SL1 PCR based library.  
Ostertagia ostertagi L3 cDNA PCR products of size >400  
nucleotides containing SL1 on the 5' end and oligo(dT) on  
the 3' end were non-directionally cloned into  
pCRII-TOPO(Invitrogen) following the Topo TA cloning  
protocol. Nematodes were provided by Dr. Louis Gasbarre  
of the USDA, Beltsville, MD (lgasbar@anri.barc.usda.gov).  
Third stage exsheathed larvae were collected from 14 day  
fecal-sphagnum moss cultures of Ostertagia eggs. The  
larvae were recovered by overnight passage on a Baermann  
apparatus, and then cleaned by passage through a 20 micron  
nylon mesh. The larvae were then subjected to a treatment  
with 1.25% chlorox to induce excystation. The larvae were  
washed with 5 changes of PBS and then pelleted and snap  
frozen in liquid nitrogen."

## ORIGIN

Alignment Scores:  
Pred. No.: 2.67e+03 Length: 396  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-606-910E-3 (1-7) x BQ626165 (1-396)

Qy 1 ProAspThrArgProAlaPro 7

Db 100 CCGGACACGAGCGCGGACCG 80

## RESULT 12

BF882322 408 bp mRNA linear EST 17-JAN-2001  
LOCUS IL3-ET0113-051200-372-F01 ET0113 Homo sapiens cDNA, mRNA sequence.  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 408)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
PUBMED 10737800  
CONTACT: Simpson A.J.G.  
COMMENT Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-ET0113-  
051200-372-F01&t3=2000-12-05&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 407.  
Location/Qualifiers  
1. .408

## FEATURES

source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="ET0113"  
/note="Organ: lung\_tumor; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

Alignment Scores:  
Pred. No.: 2.74e+03 Length: 408  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x BF882322 (1-408)

Qy 1 ProAspThrArgProAlaPro 7

Db 191 CCGGATACCCGACCGCCCA 211

## RESULT 13

CB043230/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Teladorsagia circumcincta

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
Trichostrongyloidea; Haemonchidae; Ostertaginae; Teladorsagia.  
1 (bases 1 to 411)  
Blaxter,M.L., Parkinson,J., Whittom,C., Daub,J., Guiliano,D.,  
Hall,N., Quayle,M. and Barrell,B.  
Edinburgh University/Sanger Centre Nematode EST Project  
Unpublished (2000)  
Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Aeshworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger

Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
 PCR Primers  
 FORWARD: T3  
 BACKWARD: T7PL  
 Plate: 19 row: H column: 08  
 Seq primer: SKPL  
 High quality sequence stop: 411.  
 Location/Qualifiers

## FEATURES

source  
 1..411  
 /organism="Teladorsagia circumcincta"

/mol\_type="mRNA"

/db\_xref="taxon:45464"

/clone="Tc L3 19H08"

/sex="mixed"

/dev\_stage="L3"

/clone\_lib="Teladorsagia circumcincta L3 library"

/note="Vector: Lambda Uni-ZAP; Site\_1: EcoRI; Site\_2: XhoI; Teladorsagia circumcincta is a parasitic nematode. The library was constructed from mRNA from Teladorsagia circumcincta L3 stage."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2.76e+03 Length: 411  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-606-910E-3 (1-7) x CB043230 (1-411)

QY 1 ProAspThrArgProAlaPro 7

DB 142 CCGGACAGCGGCTGCACCG 122

## RESULT 14

BU070786/c

LOCUS

DEFINITION

BU070786 453 bp mRNA linear EST 27-AUG-2002

im25c02.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6036075

5' similar to TR:Q9ULD0 Q9ULD0 KIAA1290 PROTEIN ;, mRNA sequence.

ACCESSION

BU070786

VERSION

BU070786.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,

Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,

Hillier,L., Marra,M., Pape,D., Wyllie,T., Martin,J., Blistain,A.,

Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,

Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R.,

Williams,T., Jackson,Y. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Other\_ESRs: im25c02.x1

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohph.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to

pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington

University Genome Sequencing Center For information on obtaining a

clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco

High quality sequence stop: 430.

Location/Qualifiers

source

1..453

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6036075"

/tissue\_type="insulinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="Human insulinoma"

/note="Organ: pancreas; Vector: pBluescript SK-; Site\_1:

XhoI; Site\_2: EcoRI; Constructed with lambda ZAPII system

(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to

pBluescript SK- by Dr. H. Inoue following the Washington

University protocol

(http://genome.wustl.edu/est/lambda\_protocol.shtm).

Please contact Hiroshi Inoue, MD/PhD for further

information on this library (Metabolism Division, Permutt

Laboratory, Washington University School of Medicine, Box

8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this

is a Washington University Pancreas EST project library."

US-09-606-910E-3 (1-7) x BU070786 (1-453)

## ORIGIN

Alignment Scores:  
 Pred. No.: 3.02e+03 Length: 453  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-09-606-910E-3 (1-7) x BU070786 (1-453)

QY 1 ProAspThrArgProAlaPro 7

DB 92 CCTGACACGAGCCAGCTCCG 72

## RESULT 15

CC821777/c

LOCUS

DEFINITION

CC821777 468 bp DNA linear GSS 01-JAN-2004

MBSP7C3 Uncultured virus from nearshore marine surface sediment

uncultured marine virus genomic, genomic survey sequence.

ACCESSION

CC821777

VERSION

CC821777.1

KEYWORDS

GSS.

SOURCE

uncultured marine virus

ORGANISM

uncultured marine virus

Viruses: environmental samples.

REFERENCE

AUTHORS

Breitbart,M., Felts,B., Kelley,S., Mahaffy,J.M., Nulton,J.,

Salamon,P. and Rohwer,F.

Diversity and Population Structure of a Nearshore Marine Sediment

Viral Community

Unpublished (2003)

Contact: Rohwer F

Biology Dept.

San Diego State University

5500 Campanile Dr, San Diego, CA 92102, USA

Tel: 6195941336

Fax: 619595676

Email: forest@sunstroke.sdsu.edu

Class: shotgun.

Location/Qualifiers

1..468

/organism="uncultured marine virus"

/mol\_type="genomic DNA"

/db\_xref="taxon:186617"

/clone\_lib="Uncultured virus from nearshore marine surface

sediment"

## FEATURES

source

## ORIGIN

Alignment Scores:  
 Pred. No.: 3.12e+03 Length: 468  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0

```
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%  Indels: 0
DB: 9  Gaps: 0
```

US-09-606-910E-3 (1-7) x CC821777 (1-468)

Qy 1 ProAspThrArgProAlaPro 7

Db  
145 CCAGATACCGCGCGCTCCC 125

Search completed: December 29, 2004, 14:55:20  
Job time : 2052 secs

This Page Blank (uspto)

```
Alignment Scores:
Pred. No.: 28.3 Length: 60
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-606-910E-3 (1-7) x US-10-716-293-214 (1-60)

Qy 1 ProAspThrArgProAlaPro 7
Db 1 CCAGTACAGACCGGCCCT 21

RESULT 12
US-10-471-607-3
; Sequence 3, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088857PWO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial epitope construct
US-10-471-607-3

Alignment Scores:
Pred. No.: 40.8 Length: 93
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-606-910E-3 (1-7) x US-10-471-607-3 (1-93)

Qy 1 ProAspThrArgProAlaPro 7
Db 17 CCTGACACGACCTGCCCT 37

RESULT 13
US-10-471-607-4/c
; Sequence 4, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088857PWO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificial epitope construct
US-10-471-607-4
```

```
Alignment Scores:
Pred. No.: 40.8 Length: 93
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-606-910E-3 (1-7) x US-10-471-607-4 (1-93)

Qy 1 ProAspThrArgProAlaPro 7
Db 78 CCGGACACAGCGCCGCCCA 58

RESULT 14
US-10-259-678-644
; Sequence 644, Application US/10259678
; Publication No. US20030198974A1
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Buchrieser-Brosch, Roland
; APPLICANT: Gordon, Stephen
; APPLICANT: Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/10/259,678
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/060,756
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 644
; LENGTH: 146
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-259-678-644

Alignment Scores:
Pred. No.: 59.5 Length: 146
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-606-910E-3 (1-7) x US-10-259-678-644 (1-146)

Qy 1 ProAspThrArgProAlaPro 7
Db 7 CCTGATCTCGACCGGCCCA 27

RESULT 15
US-10-471-607-5
; Sequence 5, Application US/10471607
; Publication No. US20040115740A1
; GENERAL INFORMATION:
; APPLICANT: The Victoria University of Manchester
; APPLICANT: Benson, Roderick
; TITLE OF INVENTION: Intracellular analysis.
; FILE REFERENCE: P088857PWO
; CURRENT APPLICATION NUMBER: US/10/471,607
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: GB 0108165.2
; PRIOR FILING DATE: 3001-03-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 156
; TYPE: DNA
; ORGANISM: Artificial
```

```

; FEATURE:
; OTHER INFORMATION: Artificial epitope construct
US-10-471-607-5

Alignment Scores:
Pred. No.:      62.8      Length:      156
Score:          41.00     Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     100.00% Indels:      0
DB:              17      Gaps:       0

US-09-606-910E-3 (1-7) x US-10-471-607-5 (1-156)

QY      1 ProAspThrArgProAlaPro 7
Db      17 CCTGACACGACGACCTGCCCCCT 37

Search completed: December 29, 2004, 16:16:48
Job time : 1814 secs

```



Qy 1 ProAspThrArgProAlaPro 7  
Db 517 CCGGACACCGCGCGCCCG 537  
RESULT 6  
US-08-998-416-514/C  
; Sequence 514, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippssen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgan  
; APPLICANT: Knechtle, Philipp  
; APPLICANT: Rebischung, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6239264artis Corporation  
; STREET: 3054 Cornwallis Road  
; CITY: Research Triangle Park  
; STATE: No. 6239264th Carolina  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/998,416  
; FILING DATE: 24-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: CH 0016/97  
; FILING DATE: 31-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8587  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 514:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 854 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: PAG1362UP  
; US-08-998-416-514  
Alignment Scores:  
Pred. No.: 172 Length: 854  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0  
US-09-606-910E-3 (1-7) x US-08-998-416-514 (1-854)  
Qy 1 ProAspThrArgProAlaPro 7  
Db 450 CCGGATACCGCGCCCGCCA 430  
RESULT 7  
US-10-029-517-3

; Sequence 3, Application US/10029517  
; Patent No. 6716627  
; GENERAL INFORMATION:  
; APPLICANT: Kenneth W. Dobie  
; APPLICANT: Susan J. Myers  
; TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION  
; FILE REFERENCE: RYS-0352  
; CURRENT APPLICATION NUMBER: US/10/029,517  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 107  
; SEQ ID NO 3  
; LENGTH: 1721  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (58)...(1605)  
US-10-029-517-3  
Alignment Scores:  
Pred. No.: 351 Length: 1721  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-09-606-910E-3 (1-7) x US-10-029-517-3 (1-1721)  
Qy 1 ProAspThrArgProAlaPro 7  
Db 601 CCGGACACCGCGCGCCCG 621  
RESULT 8  
US-09-252-991A-10827  
; Sequence 10827, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 10827  
; LENGTH: 1761  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-10827  
Alignment Scores:  
Pred. No.: 359 Length: 1761  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-09-606-910E-3 (1-7) x US-09-252-991A-10827 (1-1761)  
Qy 1 ProAspThrArgProAlaPro 7  
Db 1634 CCGATACCGCGCCTGCACCC 1654  
RESULT 9  
US-10-029-517-17  
; Sequence 17, Application US/10029517  
; Patent No. 6716627

GENERAL INFORMATION:  
APPLICANT: Kenneth W. Dobie  
APPLICANT: Susan J. Myers  
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION  
FILE REFERENCE: RTS-0352  
CURRENT APPLICATION NUMBER: US/10/029,517  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 107  
SEQ ID NO 17  
LENGTH: 1804  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (73)...(1500)  
US-10-029-517-17

Alignment Scores:  
Pred. No.: 368 Length: 1804  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-606-910E-3 (1-7) x US-10-029-517-17 (1-1804)  
QY 1 ProAspThrArgProAlaPro 7  
DB 496 CCGGACACCCAGCGCGCCCG 516

RESULT 10  
US-09-252-991A-10853/c  
Sequence 10853, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 10853  
LENGTH: 2256  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-10853

Alignment Scores:  
Pred. No.: 462 Length: 2256  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-606-910E-3 (1-7) x US-09-252-991A-10853 (1-2256)  
QY 1 ProAspThrArgProAlaPro 7  
DB 1529 CCGGATACCCGCGCTGCACCC 1509

RESULT 11  
US-09-620-312D-74/c  
Sequence 74, Application US/09620312D  
Patent No. 6569662  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Felyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yunging  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tillinghaast  
APPLICANT: Dmanac, Radoje T.  
TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pt\_FL\_genes Version 1.0  
SEQ ID NO 74  
LENGTH: 3718  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (90)...(3122)  
US-09-620-312D-74

Alignment Scores:  
Pred. No.: 770 Length: 3718  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-606-910E-3 (1-7) x US-09-620-312D-74 (1-3718)  
QY 1 ProAspThrArgProAlaPro 7  
DB 2919 CCTGACACCCAGCGCTCCG 2899

RESULT 12  
US-10-029-517-19  
Sequence 19, Application US/10029517  
Patent No. 6716627  
GENERAL INFORMATION:  
APPLICANT: Kenneth W. Dobie  
APPLICANT: Susan J. Myers  
TITLE OF INVENTION: ANTISENSE MODULATION OF MUCIN 1, TRANSMEMBRANE EXPRESSION  
FILE REFERENCE: RTS-0352  
CURRENT APPLICATION NUMBER: US/10/029,517  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 107  
SEQ ID NO 19  
LENGTH: 8186  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 6899  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 7155  
OTHER INFORMATION: unknown

```
; NAME/KEY: unsure
; LOCATION: 7184
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7957
; OTHER INFORMATION: unknown
; NAME/KEY: intron
; LOCATION: (2997)...(3498)
; OTHER INFORMATION: intron 1
; NAME/KEY: intron:exon junction
; LOCATION: (3498)...(3499)
; OTHER INFORMATION: intron 1:exon 2
; NAME/KEY: exon
; LOCATION: (3508)...(3599)
; OTHER INFORMATION: exon 2d
; NAME/KEY: exon:intron junction
; LOCATION: (3982)...(3983)
; OTHER INFORMATION: exon 2a:intron 2a
; NAME/KEY: intron:exon junction
; LOCATION: (4205)...(4206)
; OTHER INFORMATION: intron 2c:exon 3c
; NAME/KEY: intron:exon junction
; LOCATION: (4259)...(4260)
; OTHER INFORMATION: intron 2d:exon 3d
; NAME/KEY: exon
; LOCATION: (4260)...(4328)
; OTHER INFORMATION: exon 3d
; NAME/KEY: intron:exon junction
; LOCATION: (4632)...(4633)
; OTHER INFORMATION: intron 3:exon 4
; NAME/KEY: exon
; LOCATION: (4914)...(5035)
; OTHER INFORMATION: exon 5
; NAME/KEY: intron
; LOCATION: (5266)...(6293)
; OTHER INFORMATION: intron 6
; US-10-029-517-19

Alignment Scores:
Pred. No.: 1.72e+03 Length: 8186
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-606-910E-3 (1-7) x US-10-029-517-19 (1-8186)

Qy 1 ProAspThrArgProAlaPro 7
Db 3864 CCGGACACCGCGCGCGCCG 3884

RESULT 13
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
```

```
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 7.78e+05 Length: 4403765
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-606-910E-3 (1-7) x US-09-103-840A-2 (1-4403765)

Qy 1 ProAspThrArgProAlaPro 7
Db 375904 CCGATACCTCGACCGCCCA 375884

RESULT 14
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:
Pred. No.: 7.79e+05 Length: 4411529
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-606-910E-3 (1-7) x US-09-103-840A-1 (1-4411529)

Qy 1 ProAspThrArgProAlaPro 7
Db 375848 CCGATACCTCGACCGCCCA 375828

RESULT 15
US-09-252-991A-13024/c
; Sequence 13024, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13024
```

```
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13024

Alignment Scores:
Pred. No.: 734      Length: 1149
Score: 38.00      Matches: 6
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 85.71%      Mismatches: 0
Query Match: 92.68%      Indels: 0
DB: 4      Gaps: 0

US-09-606-910E-3 (1-7) x US-09-252-991A-13024 (1-1149)

QY 1 ProAspThrArgProAlaPro 7
Db 160 CCAGATACCCGACCTTCTCT 140
```

Search completed: December 29, 2004, 15:07:20  
Job time : 720 secs

```

PD 31-OCT-1996.
XX
PF 28-APR-1995; 95DE-01015219.
XX
XX 28-APR-1995; 95DE-01015219.
PR
XX (PECH/) PECHER G.
PA
XX Pecher G;
PI
XX WPI; 1996-486314/49.
DR
DR P-PSDB; AAW01455.
XX
XX Assay for tumour-specific cytotoxic T cells - using stimulator and target
PT B cells transfected with mucin gene sequence.
XX
XX Claim 3; Fig 1; 5pp; German.
PS
XX Stimulator and target cells are used in a novel assay to determine the
CC specific cellular immune response of a patient to tumour-associated
CC mucin. The stimulator and target cells are preferably Epstein-Barr virus-
CC immortalised chimpanzee B cells that have been transfected with a human
CC MUC1 gene sequence and treated with a glycosylation inhibitor. In
CC particular, the MUC1 gene sequence includes 15-30 tandem repeats of the
CC present 60 bp sequence
XX
SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 75.8 Length: 60
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAT44506 (1-60)

Qy 1 ProAspThrArgProAlaPro 7
Db 1 CCGGACACCGAGCGGCCCG 21

RESULT 4
AAT38664
ID AAT38664 standard; DNA; 60 BP.
XX
AC AAT38664;
XX
DT 05-AUG-1997 (first entry)
XX
DE Human MUC1 gene tandem repeat.
XX
XX Tandem nucleotide repeat; mucin; epitope; vaccine; MUC1; tumour; breast;
KW pancreas; ovary; colon; parotid gland; lung; therapy; plasmid prep4MUC1;
KW ss.
XX
XX Homo sapiens.
OS
XX
PH Key Location/Qualifiers
FT repeat_unit 1..60
FT /tag= a
FT /rpt_type= TANDEM
FT /transl_except= (pos:49..51, aa:Val)
FT /note= "Preferably 15-30 copies are present in vectors
FT which are used for transfecting irradiated, EBV-
FT immortalised autologous human B-cells"
XX
XX DE19516673-A1.
PN
XX
XX 31-OCT-1996.
PD
XX
XX 28-APR-1995; 95DE-01016673.
PF (PECH/) PECHER G.
XX
XX Pecher G;
PI
XX WPI; 1996-486530/49.
DR
DR P-PSDB; AAW04399.
XX
XX Vaccines against human, pref. mucin expressing, tumours - comprising
PT autologous B-cells transfected with mucin gene segments, treated with
XX glycosylation inhibitor.
XX
XX Claim 3; Fig 2; 5pp; German.
PS
XX B-cells are isolated from the peripheral blood of a patient and are
CC immortalised with Epstein-Barr virus (EBV) prior to transfection by
CC electroporation with a vector containing segments of the human MUC1 gene.
CC Preferably, there are 15-30 tandem repeats of the present 60 bp sequence
CC in the vector. One specific vector was designated prep4MUC1; it comprises
CC the MUC1 sequences under the control of a RSV-LTR promoter, a hygromycin
CC resistance gene under the control of a TK promoter and an EBNA-1 gene
CC sequence. Cells are incubated for 24-48 hours in culture medium
CC containing 5 mM of a glycosylation inhibitor. Finally the cells are
CC irradiated with at least 10000 rad. The cells are useful as vaccines for
CC therapy of mucin-expressing tumours, e.g. tumours of the breast,
CC pancreas, ovary, colon, parotid gland and lung
XX
SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 75.8 Length: 60
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAT38664 (1-60)

Qy 1 ProAspThrArgProAlaPro 7
Db 1 CCGGACACCGAGCGGCCCG 21

RESULT 5
AAT88166
ID AAT88166 standard; DNA; 60 BP.
XX
AC AAT88166;
XX
DT 28-APR-1998 (first entry)
XX
DE Human MUC1 tandem nucleotide repeat.
XX
XX MUC1; mucin; tandem repeat; anticancer vaccine; dendritic cell;
KW tumour associated epitope; ds.
XX
XX Homo sapiens.
OS
XX
PH Key Location/Qualifiers
FT CDS 1..60
FT /tag= a
FT /transl_except= (pos:49..51, aa:Val)
XX
XX DE19617837-A1.
PN
XX
XX 23-OCT-1997.
PD
XX
XX 19-APR-1996; 96DE-01017837.
XX
XX 19-APR-1996; 96DE-01017837.
PF
XX (PECH/) PECHER G.
XX

```

XX Pecher G;  
 XX WPI; 1997-514604/48.  
 DR P-PSDB; AAW32821.  
 XX  
 XX Anticancer vaccine containing human dendritic cells - transfected with  
 PT mucin gene containing tandem repeat sequences, and expressing tumour  
 PT associated epitope when treated with glycosylation inhibitor.  
 XX  
 XX Claim 4; Fig 2; 6pp; German.  
 XX  
 XX This sequence represents a tandem repeat found in the MUC1 gene which  
 CC encodes the glycoprotein mucin. This repeat region is used in a novel  
 CC anticancer vaccine which also contains autologous human dendritic cells,  
 CC capable of expressing tumour associated epitopes, preferably on the cell  
 CC surface, when treated with a glycosylation inhibitor. The vaccine can be  
 CC used to treat MUC1 expressing tumours, especially mammary, pancreatic,  
 CC ovarian, colon and parotid tumours  
 XX  
 SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 75.8 Length: 60  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAT88166 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 Db 1 CCGGACACCGCGCCCG 21

RESULT 6  
 AAV48320  
 ID AAV48320 standard; DNA; 60 BP.  
 XX  
 AC AAV48320;  
 XX  
 DT 20-NOV-1998 (first entry)  
 XX  
 DE Nucleotide sequence encoding MUC1 tandem repeat unit R5.  
 XX  
 KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 KW tumour; tumour-associated antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9837095-A2.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 XX New recombinant pox virus for tumour therapy - comprises DNA encoding an  
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.  
 PT  
 PS Disclosure; Page 11; 42pp; English.  
 XX  
 XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an  
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus  
 CC (RPV). The RPV was used in a pharmaceutical composition also containing  
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The  
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that  
 CC does not undergo significant genetic deletion, thereby providing an  
 CC unexpectedly stable and immunogenic pox virus. They can be used to  
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens  
 SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 75.8 Length: 60  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAV48320 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 Db 40 CCGGATACCGACCGGCCCT 60

RESULT 7  
 AAV48321  
 ID AAV48321 standard; DNA; 60 BP.  
 XX  
 AC AAV48321;  
 XX  
 DT 20-NOV-1998 (first entry)  
 XX  
 DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.  
 XX  
 KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 KW tumour; tumour-associated antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9837095-A2.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 XX New recombinant pox virus for tumour therapy - comprises DNA encoding an  
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.  
 PT  
 PS Disclosure; Page 11; 42pp; English.  
 XX  
 XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an  
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus  
 CC (RPV). The RPV was used in a pharmaceutical composition also containing  
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The  
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that  
 CC does not undergo significant genetic deletion, thereby providing an  
 CC unexpectedly stable and immunogenic pox virus. They can be used to  
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens  
 SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 75.8 Length: 60  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAV48320 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 Db 40 CCGGATACCGACCGGCCCT 60

RESULT 7  
 AAV48321  
 ID AAV48321 standard; DNA; 60 BP.  
 XX  
 AC AAV48321;  
 XX  
 DT 20-NOV-1998 (first entry)  
 XX  
 DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.  
 XX  
 KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 KW tumour; tumour-associated antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9837095-A2.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 XX New recombinant pox virus for tumour therapy - comprises DNA encoding an  
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.  
 PT  
 PS Disclosure; Page 11; 42pp; English.  
 XX  
 XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an

CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus  
 CC (RPV). The RPV was used in a pharmaceutical composition also containing  
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The  
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that  
 CC does not undergo significant genetic deletion, thereby providing an  
 CC unexpectedly stable and immunogenic pox virus. They can be used to  
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens  
 XX  
 SQ Sequence 60 BP; 9 A; 27 C; 15 G; 9 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 75.8 Length: 60  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAV48320 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 Db 40 CCGGATACCGACCGGCCCT 60

RESULT 7  
 AAV48321  
 ID AAV48321 standard; DNA; 60 BP.  
 XX  
 AC AAV48321;  
 XX  
 DT 20-NOV-1998 (first entry)  
 XX  
 DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.  
 XX  
 KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 KW tumour; tumour-associated antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9837095-A2.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 XX New recombinant pox virus for tumour therapy - comprises DNA encoding an  
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.  
 PT  
 PS Disclosure; Page 11; 42pp; English.  
 XX  
 XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an  
 CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus  
 CC (RPV). The RPV was used in a pharmaceutical composition also containing  
 CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The  
 CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that  
 CC does not undergo significant genetic deletion, thereby providing an  
 CC unexpectedly stable and immunogenic pox virus. They can be used to  
 CC prevent or treat tumours expressing MUC1 tumour-associated antigens  
 SQ Sequence 60 BP; 13 A; 27 C; 17 G; 3 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 75.8 Length: 60  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAV48320 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 Db 40 CCGGATACCGACCGGCCCT 60

RESULT 7  
 AAV48321  
 ID AAV48321 standard; DNA; 60 BP.  
 XX  
 AC AAV48321;  
 XX  
 DT 20-NOV-1998 (first entry)  
 XX  
 DE Nucleotide sequence encoding MUC1 tandem repeat unit R6.  
 XX  
 KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 KW tumour; tumour-associated antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9837095-A2.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 PD 27-AUG-1998.  
 XX  
 PF 24-FEB-1998; 98WO-US003693.  
 XX  
 PR 24-FEB-1997; 97US-0038253P.  
 XX  
 XX (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX  
 DR WPI; 1998-467492/40.  
 XX  
 XX New recombinant pox virus for tumour therapy - comprises DNA encoding an  
 PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.  
 PT  
 PS Disclosure; Page 11; 42pp; English.  
 XX  
 XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an

Alignment Scores:  
 Pred. No.: 75.8 Length: 60

Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAV48321 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 |||||  
 Db 40 CCAGACACTCGACCTGCGCCA 60

#### RESULT 8

AAV48325  
 ID AAV48325 standard; DNA; 60 BP.

XX AC AAV48325;

XX DT 20-NOV-1998 (first entry)

XX DE Nucleotide sequence encoding MUC1 tandem repeat unit R10.

XX KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 tumour; tumour-associated antigen.

XX OS Homo sapiens.

XX PN WO9837095-A2.

XX PD 27-AUG-1998.

XX PF 24-FEB-1998; 98WO-US003693.

XX PR 24-FEB-1997; 97US-0038253P.

XX PA (THER-) THERION BIOLOGICS CORP.  
 (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 (DAND ) DANA FARBER CANCER INST INC.

XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX DR WPI; 1998-467492/40.

XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an  
 immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX PS Disclosure; Page 11; 42pp; English.

XX CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an  
 immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus  
 (RPV). The RPV was used in a pharmaceutical composition also containing  
 an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The  
 recombinant pox virus therefore encodes an immunogenic MUC1 fragment that  
 does not undergo significant genetic deletion, thereby providing an  
 unexpectedly stable and immunogenic pox virus. They can be used to  
 prevent or treat tumours expressing MUC1 tumour-associated antigens

XX SQ Sequence 60 BP; 12 A; 22 C; 19 G; 7 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 75.8 Length: 60  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAV48325 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 |||||  
 Db 40 CCGATACGACCGCGCCT 60

#### RESULT 9

AAV48322  
 ID AAV48322 standard; DNA; 60 BP.

XX AC AAV48322;

XX DT 20-NOV-1998 (first entry)

XX DE Nucleotide sequence encoding MUC1 tandem repeat unit R7.

XX KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 tumour; tumour-associated antigen.

XX OS Homo sapiens.

XX PN WO9837095-A2.

XX PD 27-AUG-1998.

XX PF 24-FEB-1998; 98WO-US003693.

XX PR 24-FEB-1997; 97US-0038253P.

XX PA (THER-) THERION BIOLOGICS CORP.  
 (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 (DAND ) DANA FARBER CANCER INST INC.

XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX DR WPI; 1998-467492/40.

XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an  
 immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX PS Disclosure; Page 11; 42pp; English.

XX CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an  
 immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus  
 (RPV). The RPV was used in a pharmaceutical composition also containing  
 an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The  
 recombinant pox virus therefore encodes an immunogenic MUC1 fragment that  
 does not undergo significant genetic deletion, thereby providing an  
 unexpectedly stable and immunogenic pox virus. They can be used to  
 prevent or treat tumours expressing MUC1 tumour-associated antigens

XX SQ Sequence 60 BP; 14 A; 24 C; 16 G; 6 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 75.8 Length: 60  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-09-606-910E-3 (1-7) x AAV48322 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 |||||  
 Db 40 CCAGACACTCGACCTGCGCCA 60

#### RESULT 10

AAV48324  
 ID AAV48324 standard; DNA; 60 BP.

XX AC AAV48324;

XX DT 20-NOV-1998 (first entry)

XX DE Nucleotide sequence encoding MUC1 tandem repeat unit R9.

XX KW ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 tumour; tumour-associated antigen.

XX Homo sapiens.  
 OS WO9837095-A2.  
 PN 27-AUG-1998.  
 XX 24-FEB-1998; 98WO-US003693.  
 XX 24-FEB-1997; 97US-0038253P.  
 PR (THER-) THERION BIOLOGICS CORP.  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 XX (DAND) DANA FARBER CANCER INST INC.  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX WPI; 1998-467492/40.  
 XX New recombinant pox virus for tumour therapy - comprises DNA encoding an immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.  
 PT Disclosure; Page 11; 42pp; English.  
 XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus (RPV). The RPV was used in a pharmaceutical composition also containing an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus therefore encodes an immunogenic MUC1 fragment that does not undergo significant genetic deletion, thereby providing an unexpectedly stable and immunogenic pox virus. They can be used to prevent or treat tumours expressing MUC1 tumour-associated antigens  
 XX Sequence 60 BP; 12 A; 26 C; 14 G; 8 T; 0 U; 0 Other;  
 SQ Alignment Scores: Pred. No.: 75.8 Length: 60  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-606-910E-3 (1-7) x AAV48324 (1-60)  
 QY 1 ProAspThrArgProAlaPro 7  
 DB 40 CCGACACCGCTCCAGCTCCG 60  
 RESULT 11  
 AAV48318  
 ID AAV48318 standard; DNA; 60 BP.  
 AC AAV48318;  
 XX 20-NOV-1998 (first entry)  
 DT Nucleotide sequence encoding MUC1 tandem repeat unit R3.  
 DE ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 KW tumour; tumour-associated antigen.  
 XX Homo sapiens.  
 OS WO9837095-A2.  
 PN 27-AUG-1998.  
 XX 24-FEB-1998; 98WO-US003693.  
 XX 24-FEB-1997; 97US-0038253P.  
 PR (THER-) THERION BIOLOGICS CORP.  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 XX (DAND) DANA FARBER CANCER INST INC.  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX WPI; 1998-467492/40.  
 XX New recombinant pox virus for tumour therapy - comprises DNA encoding an immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.  
 PT Disclosure; Page 11; 42pp; English.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND) DANA FARBER CANCER INST INC.  
 XX Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX WPI; 1998-467492/40.  
 XX New recombinant pox virus for tumour therapy - comprises DNA encoding an immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.  
 PT Disclosure; Page 11; 42pp; English.  
 XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus (RPV). The RPV was used in a pharmaceutical composition also containing an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus therefore encodes an immunogenic MUC1 fragment that does not undergo significant genetic deletion, thereby providing an unexpectedly stable and immunogenic pox virus. They can be used to prevent or treat tumours expressing MUC1 tumour-associated antigens  
 XX Sequence 60 BP; 7 A; 29 C; 19 G; 5 T; 0 U; 0 Other;  
 SQ Alignment Scores: Pred. No.: 75.8 Length: 60  
 Score: 41.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-606-910E-3 (1-7) x AAV48318 (1-60)  
 QY 1 ProAspThrArgProAlaPro 7  
 DB 40 CCGACACCGCTCCAGCTCCG 60  
 RESULT 12  
 AAV48323  
 ID AAV48323 standard; DNA; 60 BP.  
 XX AAV48323;  
 XX 20-NOV-1998 (first entry)  
 DT Nucleotide sequence encoding MUC1 tandem repeat unit R8.  
 DE ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 KW tumour; tumour-associated antigen.  
 XX Homo sapiens.  
 OS WO9837095-A2.  
 PN 27-AUG-1998.  
 XX 24-FEB-1998; 98WO-US003693.  
 XX 24-FEB-1997; 97US-0038253P.  
 PR (THER-) THERION BIOLOGICS CORP.  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 XX (DAND) DANA FARBER CANCER INST INC.  
 PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;  
 XX WPI; 1998-467492/40.  
 XX New recombinant pox virus for tumour therapy - comprises DNA encoding an immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.  
 PT Disclosure; Page 11; 42pp; English.



CC The MUC1 tandem repeat units AAV48317-V48325 were used to create an immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus (RPV). The RPV was used in a pharmaceutical composition also containing an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus therefore encodes an immunogenic MUC1 fragment that does not undergo significant genetic deletion, thereby providing an unexpectedly stable and immunogenic pox virus. They can be used to prevent or treat tumours expressing MUC1 tumour-associated antigens

SQ Sequence 60 BP; 10 A; 24 C; 17 G; 9 T; 0 U; 0 Other;

Alignment Scores: 75.8 Length: 60  
 Pred. No.: 41.00 Matches: 7  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 2

US-09-606-910E-3 (1-7) x AAV48323 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 |||||  
 Db 40 CCTGACACAGCGCCGCCCA 60

RESULT 13

AAV48316  
 ID AAV48316 standard; cDNA; 60 BP.

XX AC AAV48316;

XX DT 20-NOV-1998 (first entry)

XX DE Nucleotide sequence encoding MUC1 tandem repeat unit.

XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 KW tumour; tumour-associated antigen.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT CDS 1..60

FT /\*tag= a  
 FT /product= "MUC1 tandem repeat unit"

XX PN W09837095-A2.

XX PD 27-AUG-1998.

XX PF 24-FEB-1998; 98WO-US003693.

XX PR 24-FEB-1997; 97US-0038253P.

XX PA (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.

XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX WPI; 1998-467492/40.

DR P-PSDB; AAW77229.

XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX Example 1; Page 20; 42pp; English.

XX The MUC1 tandem repeat unit was used to create an immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus (RPV). The RPV was used in a pharmaceutical composition also containing an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus therefore encodes an immunogenic MUC1 fragment that does not undergo significant genetic deletion, thereby providing an unexpectedly stable

CC and immunogenic pox virus. They can be used to prevent or treat tumours expressing MUC1 tumour-associated antigens

XX SQ Sequence 60 BP; 7 A; 33 C; 16 G; 4 T; 0 U; 0 Other;

Alignment Scores: 75.8 Length: 60  
 Pred. No.: 41.00 Matches: 7  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 2

US-09-606-910E-3 (1-7) x AAV48316 (1-60)

Qy 1 ProAspThrArgProAlaPro 7  
 |||||  
 Db 40 CCGACACAGCGCGGCCCG 60

RESULT 14

AAV48319  
 ID AAV48319 standard; DNA; 60 BP.

XX AC AAV48319;

XX DT 20-NOV-1998 (first entry)

XX DE Nucleotide sequence encoding MUC1 tandem repeat unit R4.

XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
 KW tumour; tumour-associated antigen.

XX OS Homo sapiens.

XX PN W09837095-A2.

XX PD 27-AUG-1998.

XX PF 24-FEB-1998; 98WO-US003693.

XX PR 24-FEB-1997; 97US-0038253P.

XX PA (THER-) THERION BIOLOGICS CORP.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (DAND ) DANA FARBER CANCER INST INC.

XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX WPI; 1998-467492/40.

XX PT New recombinant pox virus for tumour therapy - comprises DNA encoding an immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX Disclosure; Page 11; 42pp; English.

XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus (RPV). The RPV was used in a pharmaceutical composition also containing an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The recombinant pox virus therefore encodes an immunogenic MUC1 fragment that does not undergo significant genetic deletion, thereby providing an unexpectedly stable and immunogenic pox virus. They can be used to prevent or treat tumours expressing MUC1 tumour-associated antigens

XX SQ Sequence 60 BP; 10 A; 23 C; 13 G; 14 T; 0 U; 0 Other;

Alignment Scores: 75.8 Length: 60  
 Pred. No.: 41.00 Matches: 7  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 2

US-09-606-910E-3 (1-7) x AAV48319 (1-60)

QY 1 ProAspThrArgProAlaPro 7  
|||  
DB 40 CCAGATACTCGCCGACCTCCA 60

## RESULT 15

AAV48317  
ID AAV48317 standard; DNA; 60 BP.

XX  
AC AAV48317;

XX  
DT 20-NOV-1998 (first entry)

DE Nucleotide sequence encoding MUC1 tandem repeat unit R2.

XX ss; MUC1; recombinant pox virus; cytotoxic T-lymphocyte; immunogen;  
KW tumour; tumour-associated antigen.

XX OS Homo sapiens.

XX PN WO9837095-A2.

XX PD 27-AUG-1998.

XX PF 24-FEB-1998; 98WO-US003693.

XX PR 24-FEB-1997; 97US-0038253P.

XX (THER-) THERION BIOLOGICS CORP.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA (DAND ) DANA FARBEN CANCER INST INC.

XX PI Schlom J, Kantor J, Kufe D, Panicali D, Gritz L;

XX WPI; 1998-467492/40.

XX New recombinant pox virus for tumour therapy - comprises DNA encoding an  
PT immunogenic mini-MUC1 fragment comprising 5-25 MUC1 tandem repeat units.

XX Disclosure; Page 11; 42pp; English.

XX The MUC1 tandem repeat units AAV48317-V48325 were used to create an  
CC immunogenic mini-MUC1 fragment for inclusion in a recombinant pox virus  
CC (RPV). The RPV was used in a pharmaceutical composition also containing  
CC an immunomodulator to generate MUC1 specific cytotoxic T-lymphocytes. The  
CC recombinant pox virus therefore encodes an immunogenic MUC1 fragment that  
CC does not undergo significant genetic deletion, thereby providing an  
CC unexpectedly stable and immunogenic pox virus. They can be used to  
CC prevent or treat tumours expressing MUC1 tumour-associated antigens

XX SQ Sequence 60 BP; 17 A; 21 C; 13 G; 9 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	75.8	Length:	60
Score:	41.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-606-910E-3 (1-7) x AAV48317 (1-60)

QY 1 ProAspThrArgProAlaPro 7

DB 40 CCTGATACAAAGACCTGCACCT 60

Search completed: December 29, 2004, 13:38:36  
Job time : 334 secs

```

SOURCE      unidentified
ORGANISM     unclassified.
REFERENCE    1 (bases 1 to 60)
AUTHORS      Xing,Px. and McKenzie,I.F.C.
TITLE        Monoclonal antibodies specific for human polymorphic epithelial
              mucin
JOURNAL      Patent: EP 0369816-A 27 23-MAY-1990;
              THE UNIVERSITY OF MELBOURNE
FEATURES     Location/Qualifiers
              source
                1..60
                /organism="unidentified"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN
Alignment Scores:
Pred. No.:      70.7      Length:      60
Score:          41.00     Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels:      0
DB:             6        Gaps:       0
US-09-606-910E-3 (1-7) x A09845 (1-60)
Qy            1 ProAspThrArgProAlaPro 7
Db            1 CCGGACACCGCGCGGCCCG 21
RESULT 6
AXI33974
LOCUS        AXI33974      60 bp      DNA      linear      PAT 15-MAY-2001
DEFINITION   Sequence 1 from Patent WO0124832.
ACCESSION    AXI33974
VERSION      AXI33974.1 GI:14139915
KEYWORDS     synthetic construct
              synthetic construct
              artificial sequences.
REFERENCE    1
AUTHORS      Pecher,G.
TITLE        Pharmaceutical composition for treating and preventing human
              tumors, which express the tumor antigen mucin and/or the
              carcinoembryonic antigen (ce a), and the use thereof
JOURNAL      Patent: WO 0124832-A 1 12-APR-2001;
              Pecher, Gabriele (DE)
FEATURES     Location/Qualifiers
              source
                1..60
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
              <1..>60
              /note="unnamed protein product"
              /codon_start=1
              /transl_table=11
              /protein_id="CAC38972.1"
              /db_xref="GI:14139916"
              /translation="PDTRPAPGSTAPPAGHGTSA"
CDS
Alignment Scores:
Pred. No.:      70.7      Length:      60
Score:          41.00     Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels:      0
DB:             6        Gaps:       0
US-09-606-910E-3 (1-7) x AXI33974 (1-60)
Qy            1 ProAspThrArgProAlaPro 7

```

```

Db            1 CCGGACACCGCGCGGCCCG 21
RESULT 7
HUMMUCIN
LOCUS        HUMMUCIN      60 bp      mRNA      linear      PRI 27-JUL-1994
DEFINITION   Human epithelial mucin tandem repeat sequence of the partial cds.
ACCESSION    M26316
VERSION      M26316.1 GI:516622
KEYWORDS     mucin.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 60)
AUTHORS      Xing,P.X., Tjandra,J.J., Reynolds,K., McLaughlin,P.J., Purcell,D.F.
              and McKenzie,I.F.
TITLE        Reactivity of anti-human milk fat globule antibodies with synthetic
              peptides
JOURNAL      J. Immunol. 142 (10), 3503-3509 (1989)
MEDLINE      89235154
PUBMED       2715633
COMMENT      On Jul 28, 1994 this sequence version replaced gi:341467.
              Original source text: Homo sapiens cDNA to mRNA.
FEATURES     Location/Qualifiers
              source
                1..60
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
              <1..>60
              /note="epithelial mucin tandem repeat sequence"
              /codon_start=1
              /product="mucin"
              /protein_id="AAA36336.1"
              /db_xref="GI:516623"
              /translation="PDTRPAPGSTAPPAGHGTSA"
ORIGIN
Alignment Scores:
Pred. No.:      70.7      Length:      60
Score:          41.00     Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels:      0
DB:             9        Gaps:       0
US-09-606-910E-3 (1-7) x HUMMUCIN (1-60)
Qy            1 ProAspThrArgProAlaPro 7
Db            1 CCGGACACCGCGCGGCCCG 21
RESULT 8
AXI80273
LOCUS        AXI80273      86 bp      DNA      linear      PAT 06-AUG-2001
DEFINITION   Sequence 31 from Patent WO0146228.
ACCESSION    AXI80273
VERSION      AXI80273.1 GI:15132244
KEYWORDS     synthetic construct
              synthetic construct
              artificial sequences.
REFERENCE    1
AUTHORS      Crowe,J.S. and Ellis,J.H.
TITLE        Nucleic acid vaccination
JOURNAL      Patent: WO 0146228-A 31 28-JUN-2001;
              GLAXO GROUP LIMITED (GB)
FEATURES     Location/Qualifiers
              source
                1..86
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Oligonucleotide"
ORIGIN

```

Alignment Scores: 97.7 Length: 86  
Pred. No.: 41.00 Matches: 7  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6

US-09-606-910E-3 (1-7) x AX180273 (1-86)

Qy 1 ProAspThrArgProAlaPro 7  
Db 33 CCGGACACGAGCGGCCCG 53

RESULT 9  
AX180274/c AX180274 92 bp DNA linear PAT 06-AUG-2001  
LOCUS Sequence 32 from Patent WO0146228.  
DEFINITION AX180274  
ACCESSION AX180274  
VERSION AX180274.1 GI:15132245  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1  
AUTHORS Crowe,J.S. and Ellis,J.H.  
TITLE Nucleic acid vaccination  
JOURNAL Patent: WO 0146228-A 32 28-JUN-2001;  
GLAXO GROUP LIMITED (GB)  
FEATURES Location/Qualifiers  
source  
1..92  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide"

ORIGIN  
Alignment Scores: 104 Length: 92  
Pred. No.: 41.00 Matches: 7  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6

US-09-606-910E-3 (1-7) x AX180274 (1-92)

Qy 1 ProAspThrArgProAlaPro 7  
Db 60 CCGGACACGAGCGGCCCG 40

RESULT 10  
AX703425  
LOCUS AX703425 93 bp DNA linear PAT 03-APR-2003  
DEFINITION Sequence 3 from Patent WO02086505.  
ACCESSION AX703425  
VERSION AX703425.1 GI:29538416  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1  
AUTHORS Benson,R.S.  
TITLE Intracellular analysis  
JOURNAL Patent: WO 02086505-A 3 31-OCT-2002;  
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)  
FEATURES Location/Qualifiers  
source  
1..93  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Artificial epitope construct"

## ORIGIN

Alignment Scores: 105 Length: 93  
Pred. No.: 41.00 Matches: 7  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6

US-09-606-910E-3 (1-7) x AX703425 (1-93)

Qy 1 ProAspThrArgProAlaPro 7  
Db 17 CCGGACACGAGCGGCCCT 37

RESULT 11  
AX703426/c AX703426 93 bp DNA linear PAT 03-APR-2003  
LOCUS Sequence 4 from Patent WO02086505.  
DEFINITION AX703426  
ACCESSION AX703426  
VERSION AX703426.1 GI:29538417  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1  
AUTHORS Benson,R.S.  
TITLE Intracellular analysis  
JOURNAL Patent: WO 02086505-A 4 31-OCT-2002;  
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)  
FEATURES Location/Qualifiers  
source  
1..93  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Artificial epitope construct"

## ORIGIN

Alignment Scores: 105 Length: 93  
Pred. No.: 41.00 Matches: 7  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6

US-09-606-910E-3 (1-7) x AX703426 (1-93)

Qy 1 ProAspThrArgProAlaPro 7  
Db 78 CCGGACACGAGCGGCCCA 58

RESULT 12  
AX192396/c AX192396 120 bp DNA linear PAT 15-AUG-2001  
LOCUS Sequence 2 from Patent EP1103623.  
DEFINITION AX192396  
ACCESSION AX192396  
VERSION AX192396.1 GI:15210363  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Taylor-Papadimitriou,J., Burchell,J. and Gendler,S.  
TITLE Human mucin core protein: nucleic acid probes, peptide fragments  
and antibodies thereto, and uses thereof in diagnostic and  
therapeutic methods  
JOURNAL Patent: EP 1103623-A 2 30-MAY-2001;  
IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)  
FEATURES Location/Qualifiers  
source  
1..120

```
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 132 Length: 120
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-606-910E-3 (1-7) x AX192396 (1-120)

Qy 1 ProAspThrArgProAlaPro 7
Db 108 CCGGACACCGCGCGCCCG 88

RESULT 13
LOCUS AX192397 120 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 3 from Patent EP1103623.
ACCESSION AX192397
VERSION AX192397.1 GI:15210364
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Taylor-Papadimitriou, J., Burchell, J. and Gendler, S.
TITLE Human mucin core protein: nucleic acid probes, peptide fragments
and antibodies thereto, and uses thereof in diagnostic and
therapeutic methods
JOURNAL Patent: EP 1103623-A 3 30-MAY-2001;
IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED (GB)
FEATURES
source
1. .120
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 132 Length: 120
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-606-910E-3 (1-7) x AX192397 (1-120)

Qy 1 ProAspThrArgProAlaPro 7
Db 13 CCGGACACCGCGCGCCCG 33

RESULT 14
LOCUS BD000571/c 120 bp DNA linear PAT 31-JAN-2002
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein.
ACCESSION BD000571
VERSION BD000571.1 GI:18623684
KEYWORDS JP 2000333675-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Papadimitrov, J.T., Jendora, S. and Bachieru, J.
TITLE Human polymorphic epithelial mucin core protein and nucleic acid
```

```
encoding the protein
Patent: JP 2000333675-A 2 05-DEC-2000;
IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
PN JP 2000333675-A/2
PD 05-DEC-2000
PF 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269, 07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306, 09-NOV-1987 GB 8726172 PI JOYCE
TAYLOR PAPADIMITROV, SANDRA JENDORA, JOY BACHIERU PC
C12N15/02, A61K38/00, A61K39/395, A61K49/00, A61P35/00, PC
C07K14/47,
PC C07K16/44, C12N5/10, C12P21/08// (C12N15/02, C12R1:91), (C12N5/10,
C12R1:91),
PC C12N15/00, A61K37/02, C12N5/00, (C12N15/00, C12R1:91), (C12N5/00,
C12R1:91)
CC
CC FH Key Location/Qualifiers
FT source 1. .120
FT /organism="Homo sapiens (human)"

FEATURES
source
1. .120
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 132 Length: 120
Score: 41.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-606-910E-3 (1-7) x BD000571 (1-120)

Qy 1 ProAspThrArgProAlaPro 7
Db 108 CCGGACACCGCGCGCCCG 88

RESULT 15
LOCUS BD000572 120 bp DNA linear PAT 31-JAN-2002
DEFINITION Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein.
ACCESSION BD000572
VERSION BD000572.1 GI:18623685
KEYWORDS JP 2000333675-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Papadimitrov, J.T., Jendora, S. and Bachieru, J.
TITLE Human polymorphic epithelial mucin core protein and nucleic acid
encoding the protein
Patent: JP 2000333675-A 3 05-DEC-2000;
IMPERIAL CANCER RESEARCH TECHNOLOGY LTD
OS Homo sapiens (human)
PN JP 2000333675-A/3
PD 05-DEC-2000
PF 26-APR-2000 JP 2000125724
PR 07-JAN-1987 GB 8700269, 07-JAN-1987 GB 8700279 PR
22-APR-1987 US 041306, 09-NOV-1987 GB 8726172 PI JOYCE
TAYLOR PAPADIMITROV, SANDRA JENDORA, JOY BACHIERU PC
C12N15/02, A61K38/00, A61K39/395, A61K49/00, A61P35/00, PC
C07K14/47,
PC C07K16/44, C12N5/10, C12P21/08// (C12N15/02, C12R1:91), (C12N5/10,
C12R1:91),
PC C12N15/00, A61K37/02, C12N5/00, (C12N15/00, C12R1:91), (C12N5/00,
C12R1:91)
CC
```

PH Key Location/Qualifiers  
FT source 1. .120 /organism='Homo sapiens (human)'

FEATURES  
source Location/Qualifiers  
1. .120 /organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

ORIGIN

Alignment Scores:  
Pred. No.: 132 Length: 120  
Score: 41.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-606-910E-3 (1-7) x BD000572 (1-120)

Qy 1 ProAspThrArgProAlaPro 7  
Db 13 CCGGACACCGAGCGGCCCG 33

Search completed: December 29, 2004, 14:21:05  
Job time : 2544 secs

RT "High density O-glycosylation on tandem repeat peptide from secretory  
MUC1 of T47D breast cancer cells.";  
J. Biol. Chem. 274:18165-18172(1999).  
[20]  
RP POLYMORPHISM WITHIN THE REPEAT.  
RX MEDLINE=21335966; PubMed=11350974; DOI=10.1074/jbc.M103187200;  
Engelmann K., Balduz S.E., Hanisch F.-G.;  
RA "Identification and topology of variant sequences within individual  
repeat domains of the human epithelial tumor mucin MUC1.";  
J. Biol. Chem. 276:27764-27769(2001).  
[21]  
RP CHARACTERIZATION OF ISOFORM Y, AND MUTAGENESIS OF ASP-1116.  
RX MEDLINE=92211485; PubMed=10197628;  
Baruch A., Hartmann M.-L., Yoeli M., Adereth Y., Greenstein S.,  
RA Stadler Y., Skornik Y., Zaretsky J., Smorodinsky N.I., Keydar I.,  
Wreschner D.H.;  
RA "The breast cancer-associated MUC1 gene generates both a receptor and  
its cognate binding protein.";  
Cancer Res. 59:1552-1561(1999).  
[22]  
RP PARTIAL SEQUENCE, AND CHARACTERIZATION OF CLEAVAGE SITE.  
RX MEDLINE=21240104; PubMed=11341784; DOI=10.1006/bbrc.2001.4775;  
Parry S., Silverman H.S., McDermott K., Willis A., Hollingsworth M.A.,  
RA Harris A.;  
RA "Identification of MUC1 proteolytic cleavage sites in vivo.";  
Biochem. Biophys. Res. Commun. 283:715-720(2001).  
[23]  
RP CHARACTERIZATION.  
RX MEDLINE=21836452; PubMed=11847293;  
Wreschner D.H., McGuckin M.A., Williams S.J., Baruch A., Yoeli M.,  
RA Ziv R., Okun L., Zaretsky J., Smorodinsky N., Keydar I., Neophytou P.,  
RA Stacey M., Lin H.-H., Gordon S.;  
RA "Generation of ligand-receptor alliances by 'SEA' module-mediated  
cleavage of membrane-associated mucin proteins.";  
Protein Sci. 11:698-706(2002).  
[24]  
RP PHOSPHORYLATION.  
Query Match 100.0%; Score 41; DB 1; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 PDTRPAP 7  
Db 142 PDTRPAP 148  
RESULT 3  
Q9V052 PRELIMINARY; PRT; 514 AA.  
AC Q9V052  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Major extracellular endo-1,4-beta-glucanase (Cellulase).  
GN Name=celB-like; OrderedLocusNames=PYRAB09400; ORFNames=PAB0632;  
OS Pyrococcus abyssi.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=29292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GES / Orsay;  
RX MEDLINE=22511545; PubMed=12622808;  
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,  
RA Poch O., Frier D., Querellou J., Ripp R., Thierry J.-C.,  
RA "An integrated analysis of the genome of the hyperthermophilic  
archaeon Pyrococcus abyssi.";  
Mol. Microbiol. 47:1495-1512(2003).  
RL EMBL; AJ248285; CAB49854.1; -.  
DR PIR; E75142; E75142.  
DR HSSP; P54583; 1ECE.

DR GO: 0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.  
DR GO: 0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR001547; Glyco\_hydro\_5.  
DR Pfam; PF00150; Cellulase; 1.  
DR PROSITE; PS00659; GLYCOSYL\_HVDROL\_F5; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 514 AA; 59980 MW; CF80CAB1525337C5 CRC64;  
Query Match 92.7%; Score 38; DB 2; Length 514;  
Best Local Similarity 85.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 PDTRPAP 7  
Db 101 PDTRPAP 107  
RESULT 4  
O18977 PRELIMINARY; PRT; 4135 AA.  
AC O18977  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Tenascin-X.  
GN Name=TN-X;  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=97426436; PubMed=9278449;  
RA Eleftheriou F., Exposito J.Y., Garrone R., Lethias C.;  
RT "Characterization of flexilin, the bovine tenascin-X.";  
J. Biol. Chem. 272:22866-22874(1997).  
CC -1- SIMILARITY: Contains 19 EGF-like domains.  
DR EMBL; Y11915; CAA72671.1; -.  
DR PIR; T42629; T42629.  
DR HSSP; P05107; IL3Y.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR InterPro; IPR003961; FN III.  
DR InterPro; IPR008957; FN III-like.  
DR Pfam; PF00008; EGF; 11.  
DR Pfam; PF00147; Fibrinogen\_C; 1.  
DR Pfam; PF00041; fn3; 30.  
DR SMART; SM00181; EGF; 9.  
DR SMART; SM00186; FBG; 1.  
DR SMART; SM00060; FN3; 30.  
DR PROSITE; PS00022; EGF\_1; 19.  
DR PROSITE; PS01186; EGF\_2; 19.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.  
DR PROSITE; PS00853; FN3; 30.  
KW EGF-like domain.  
SQ SEQUENCE 4135 AA; 447381 MW; D903653C3444F4A9 CRC64;  
Query Match 92.7%; Score 38; DB 2; Length 4135;  
Best Local Similarity 85.7%; Pred. No. 1.4e+03;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 PDTRPAP 7  
Db 3240 PDTRPAP 3246  
RESULT 5  
PCLO\_CHICK  
ID PCLO\_CHICK STANDARD; PRT; 5120 AA.  
AC Q9PU36;

DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Piccolo protein (Aczonin) (Fragment).  
GN Name=PCLO; Synonyms=ACZ;  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99439764; PubMed=10508862;  
RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,  
RA Killmann M.W.;  
RT "Aczonin, a 550-kd putative scaffolding protein of presynaptic active  
RT zones, shares homology regions with rim and bassoon and binds  
RT profilin.";  
RL J. Cell Biol. 147:151-162(1999).  
CC -!- FUNCTION: May act as a scaffolding protein involved in the  
CC organization of synaptic active zones and in synaptic vesicle  
CC trafficking (By similarity).  
CC -!- SUBUNIT: Interacts with Rabac1/Pral and profilin (By similarity).  
CC -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of  
CC synaptic junctions (By similarity).  
CC -!- DOMAIN: C2 domain 1 is involved in binding calcium and  
CC phospholipids. Calcium binds with low affinity but with high  
CC specificity and induces a large conformational change.  
CC -!- SIMILARITY: Contains 2 C2 domains.  
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; Y19187; CAB60725.1; -.  
CC HSP; P04410; IAZ5.  
CC GO; GO:0045202; C:synapse; ISS.  
CC GO; GO:0005509; F:calcium ion binding; ISS.  
CC GO; GO:0005544; F:calcium-dependent phospholipid binding; ISS.  
CC GO; GO:0005522; F:profilin binding; ISS.  
CC GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.  
CC GO; GO:0016080; P:synaptic vesicle targeting; ISS.  
CC InterPro; IPR000008; C2.  
CC InterPro; IPR001478; PDZ.  
CC InterPro; IPR001565; Synaptotagmin.  
CC InterPro; IPR008899; Znf\_piccolo.  
CC Pfam; PF00168; C2; 2.  
CC Pfam; PF00595; PDZ; 1.  
CC Pfam; PF05715; zf-piccolo; 2.  
CC PRINTS; PR00399; SYNAPTOTAGMN.  
CC SMART; SM00239; C2; 2.  
CC SMART; SM00228; PDZ; 1.  
CC PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
CC PROSITE; PS50004; C2\_DOMAIN\_2; 2.  
CC PROSITE; PS50106; PDZ; 1.  
CC Calcium/phospholipid-binding; Metal-binding; Repeat; Zinc;  
KW Zinc-finger.  
FT NON\_TER 1  
FT DOMAIN 258 357 10 X 10 AA tandem approximate repeats of  
FT P-A-K-P-Q-P-Q-P-X.  
FT ZN\_FING 368 392 C4-type (Potential).  
FT ZN\_FING 836 859 C4-type (Potential).  
FT DOMAIN 2324 2343 Poly-Pro.  
FT DOMAIN 4414 4493 PDZ.  
FT DOMAIN 4627 4726 C2 domain 1.  
FT DOMAIN 5003 5094 C2 domain 2.  
FT SEQUENCE 5120 AA; 560751 MW; A658D9891B65B412 CRC64;

Query Match 92.7%; Score 38; DB 1; Length 5120;  
Best Local Similarity 85.7%; Pred. No. 1.8e+03;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PDTRPAP 7  
DB 643 PDTRPAP 649  
RESULT 6  
Q84559 PRELIMINARY; PRT; 111 AA.  
AC Q84559;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE A239L protein.  
GN Name=A239L;  
OS Paramesidium bursaria Chlorella virus 1 (PBCV-1).  
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.  
OX NCBI\_TaxID=10506;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96187795; PubMed=8614977;  
RA Lu Z., Li Y., Que Q., Kutish G.F., Rock D.L., Van Etten J.L.;  
RT "Analysis of 94 kb of the chlorella virus PBCV-1 330-kb genome: map  
RT positions 88 to 182";  
RL Virology 216:102-123(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20013326; PubMed=10544099;  
RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,  
RA Lisee A.D., Nickerson K.W., Van Etten J.L.;  
RT "Chlorella virus PBCV-1 encodes a functional homosperridine  
RT synthase.";  
RL Virology 263:254-262(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20478054; PubMed=11021991;  
RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;  
RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus  
RT PBCV-1.";  
RL Virology 276:27-36(2000).  
DR EMBL; U42580; AAC96607.1; -.  
DR PIR; T17730; T17730.  
SQ SEQUENCE 111 AA; 12721 MW; 1CAC3DF047017E8 CRC64;  
Query Match 90.2%; Score 37; DB 2; Length 111;  
Best Local Similarity 85.7%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PDTRPAP 7  
DB 13 PDTRPVP 19  
RESULT 7  
Q6J2G2 PRELIMINARY; PRT; 186 AA.  
AC Q6J2G2;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Single-stranded binding protein.  
GN ORFNames=PMA4326A42;  
OS Pseudomonas syringae (pv. maculicola).  
OC Plasmid pPMA4326A.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=59511;  
RN [1]  
RP SEQUENCE FROM N.A.



```

RC STRAIN=ES4326;
RX PubMed=15262947;
RA Stavrinides J., Guttman D.S.;
RT "Nucleotide Sequence and Evolution of the Five-Plasmid Complement of
the Phytopathogen Pseudomonas syringae pv. maculicola ES4326.";
RL J. Bacteriol 186:5101-5115(2004).
CC -1- SIMILARITY: Contains 1 SSB domain.
DR EMBL; AY603979; AAT35159.1; -.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR011344; SSB.
DR InterPro; IPR000424; SSB/primosomal_n.
DR InterPro; IPR010913; SS_binding.
DR Pfam; PF00436; SSB; 1.
DR TIGRFAMs; TIGR00621; ssb; 1.
DR PROSITE; PS0935; SSB; 1.
KW DNA replication; DNA-binding; Plasmid.
SQ SEQUENCE 186 AA; 20608 MW; DF790B4F6FB20B2A CRC64;

Query Match 90.2%; Score 37; DB 2; Length 186;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
||:||||
Db 160 PDSRPAP 166

RESULT 8
AAT35159 PRELIMINARY; PRT; 186 AA.
ID AAT35159
AC AAT35159;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Single-stranded binding protein.
GN PMA4326A42.
OS Pseudomonas syringae (pv. maculicola).
OG Plasmid pPMA4326A.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas; Pseudomonas syringae group genomosp. 3.
OC NCBI_TaxID=59511;
RN [1]_TaxID=59511;
RP SEQUENCE FROM N.A.
RC STRAIN=ES4326;
RA Stavrinides J., Guttman D.S.;
RT "Nucleotide Sequence and Evolution of the Five-Plasmid Complement of
the Phytopathogen Pseudomonas syringae pv. maculicola ES4326.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY603979; AAT35159.1; -.
KW Plasmid.
SQ SEQUENCE 186 AA; 20608 MW; DF790B4F6FB20B2A CRC64;

Query Match 90.2%; Score 37; DB 2; Length 186;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
||:||||
Db 160 PDSRPAP 166

RESULT 9
Q8QZS5 PRELIMINARY; PRT; 190 AA.
ID Q8QZS5
AC Q8QZS5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE RIKEN cDNA 5830442J12 (Mus musculus adult male thymus cDNA, RIKEN
full-length enriched library, clone:5830442J12 product:hypothetical
protein, full insert sequence).
DE Name=5830442J12RIK;
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=Mix FVB/N;
RT TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanches A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=Mix FVB/N;
RT TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [4]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [6]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

```

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayaehizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [8]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Thymus;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsumura T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC035512; AAH25512.1; --  
DR EMBL; AK017982; BAC25537.1; --  
DR MGD; MGI:1923322; 5830442J12Rik.  
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.  
DR GO; GO:0004842; P:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0008270; P:zinc ion binding; IEA.  
DR GO; GO:0016567; P:protein ubiquitination; IEA.  
DR InterPro; IPR001841; Zn\_finger.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS00089; ZF\_RING\_2; 1.  
KW Hypothetical protein, Metal-binding; Zinc; Zinc-finger.  
SQ SEQUENCE 190 AA; 21625 MW; 0671B05D56DB7FB1 CRC64;  
  
Query Match 90.2%; Score 37; DB 2; Length 190;  
Best Local Similarity 85.7%; Pred. No. 88;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 PDTRPAP 7  
DB 136 PDTRPVP 142  
  
RESULT 10  
Q7QF20  
ID Q7QF20 PRELIMINARY; PRT; 249 AA.  
AC Q7QF20;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE AGCP13330 (Fragment)  
GN Name=agCG54728; ORFNames=ENSANGG00000009896;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
OX NCBI\_TaxID=180454;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PEST;  
RA Anopheles Genome Sequencing Consortium;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAAB01008846; EAA06446.1; --  
DR GO; GO:0005489; P:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR002109; Glutaredoxin.  
DR Pfam; PF00462; Glutaredoxin; 1.  
FT NON\_TER  
SQ SEQUENCE 249 AA; 28152 MW; F0610FB5E3379629 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 249;  
Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 PDTRPAP 7  
DB 28 PDTRPGP 34  
  
RESULT 11  
Q82QQ1  
ID Q82QQ1 PRELIMINARY; PRT; 384 AA.  
AC Q82QQ1;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=SAV454;  
OS Streptomyces avermitilis.  
OC Bacteria; Actinobacteria; Actinomycetales;  
OC Streptomycinae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=33903;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MA-4680;  
RX MEDLINE=21477403; PubMed=11572948;  
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,  
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,  
RT "Genome sequence of an industrial microorganism Streptomyces  
RT avermitilis: deducing the ability of producing secondary  
RT metabolites";  
RN [2]  
RP Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
RP SEQUENCE FROM N.A.  
RC STRAIN=MA-4680;  
RX MEDLINE=22608306; PubMed=12692562;  
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
RA Sakaki Y., Hattori M., Omura S.,  
RT "Complete genome sequence and comparative analysis of the industrial  
RT microorganism Streptomyces avermitilis.";  
RL Nat. Biotechnol. 21:528-531(2003).  
DR InterPro; IPR00169; Pept\_cys\_acsite.  
DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; UNKNOWN\_1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 384 AA; 42158 MW; 2AB4A385EAD71046 CRC64;  
  
Query Match 90.2%; Score 37; DB 2; Length 384;  
Best Local Similarity 85.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 PDTRPAP 7  
DB 163 PDTRPGP 169  
  
RESULT 12  
Q88LE1  
ID Q88LE1 PRELIMINARY; PRT; 911 AA.  
AC Q88LE1;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=PP1993;  
OS Pseudomonas putida (strain KT2440).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=160488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22423060; PubMed=12534463;

```

RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzaple E.K., Scanlan D., Tran K.,
RA Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016781; AAC67608.1; -.
DR TIGR; PP1993; -.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR002482; LysM.
DR Pfam; PF01476; LysM; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
KW Complete proteome, Hypothetical protein.
SQ SEQUENCE 911 AA; 96918 MW; F6EC7861AECFA2F4 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 911;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
Db 366 PDTQPAP 372

RESULT 13
O15870
ID O15870 PRELIMINARY; PRT; 2706 AA.
AC O15870;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PfEMP1 (Fragment).
GN Name=R29R+var1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IT 4/25/5; PubMed=9230440;
RX MEDLINE=97373957;
RA Rowe J.A., Moulds J.M., Newbold C.I., Miller L.H.;
RT "Plasmodium falciparum rosetting is mediated by PfEMP1 and requires complement receptor 1.";
RL Nature 388:292-295(1997).
DR EMBL; Y13402; CAA73831.1; -.
DR EMBL; Y13403; CAA73831.1; JOINED.
DR PIR; T28155; T28155.
DR GO; GO:0005539; F:glycosaminoglycan binding; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP; 2.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
FT NON_TER 2706
SQ SEQUENCE 2706 AA; 308164 MW; 1C33D55AD317D68 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 2706;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
Db 2450 PDTRPVP 2456

```

```

RESULT 14
Q84187
ID Q84187 PRELIMINARY; PRT; 125 AA.
AC Q84187;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86200383; PubMed=3754588;
RA Bellini W.J., Englund G., Richardson C.D., Rozenblatt S.,
RA Lazarini R.A.;
RT "Matrix genes of measles virus and canine distemper virus: cloning,
RT nucleotide sequences, and deduced amino acid sequences.";
RL J. Virol. 58:408-416(1986).
DR EMBL; M12668; AAA66617.1; -.
KW Hypothetical protein.
SQ SEQUENCE 125 AA; 13356 MW; 4C9482BA3D7A06F6 CRC64;

Query Match 87.8%; Score 36; DB 2; Length 125;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
Db 49 PDTRPVP 55

```

```

RESULT 15
Q8PLY4
ID Q8PLY4 PRELIMINARY; PRT; 198 AA.
AC Q8PLY4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TonB-like protein.
GN OrderedLocustNames=XAC1651;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=1204217;
RA da Silva A.C.R., Ferri J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferri M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos S.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE011797; AAM36519.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. .; IEA.
DR GO; GO:0008565; P:protein transporter activity; IEA.
DR GO; GO:0015031; P:protein transport; IEA.

```

DR InterPro; IPR006260; TonB\_C.  
 DR TIGRFAMs; TIGR01352; tonB\_Cterm; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 198 AA; 21748 MW; FB9DF173EDC52961 CRC64;  
 Query Match 87.8%; Score 36; DB 2; Length 198;  
 Best Local Similarity 85.7%; Pred. NO. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 PDTRPAP 7  
 Db |||||  
 70 PDTRPLP 76

Search completed: December 29, 2004, 13:11:13  
 Job time : 97 secs

```
Db      101 PDTRPSP 107

RESULT 4
T42629
N:Alternate names: flexilin
C:Species: Bos primigenius taurus (cattle)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42629
R:Eleftheriou, F.; Exposito, J.Y.; Garrone, R.; Lethias, C.
J. Biol. Chem. 272, 22866-22874, 1997
A:Title: Characterization of the bovine tenascin-X.
A:Reference number: 222180; MUID:97426436; PMID:9278449
A:Accession: T42629
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-4135 <ELE>
A:Cross-references: UNIPROT:O18977; EMBL:Y11915; NID:G2462978; PIDN:CAA72671.1; PID:G24629
C:Genetics:
C:Superfamily: tenascin-X; BGF homology; fibrinogen beta/gamma homology; fibronectin type
C:Keywords: extracellular matrix; glycoprotein; heptad repeat

Query Match      92.7%; Score 38; DB 2; Length 4135;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PDTRPAP 7
      |||||
Db      3240 PDTRPAP 3246

RESULT 5
T17730
hypothetical protein A239L - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T17730
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T17730
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-111 <GRA>
A:Cross-references: UNIPROT:Q84559; EMBL:U42580; NID:G4028896; PIDN:AAC96607.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A:Note: A239L
C:Superfamily: Chlorella virus PBCV-1 hypothetical protein A239L

Query Match      90.2%; Score 37; DB 2; Length 111;
Best Local Similarity 85.7%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 PDTRPAP 7
      |||||
Db      13 PDTRPVP 19

RESULT 6
T28155
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum) (fragments)
N:Alternate names: erythrocyte membrane binding protein 1 (EMBP1)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28155
R:Rowe, J.A.; Moulds, J.M.; Newbold, C.I.; Miller, L.H.
Nature 388, 292-295, 1997
A:Title: Plasmodium falciparum rosetting is mediated by PfEMP1 and requires complement
A:Reference number: Z20477; MUID:97373957; PMID:9230440
A:Accession: T28155

A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-2706 <ROW>
A:Cross-references: UNIPROT:O15870; EMBL:Y13402; PIDN:CAA73831.1
A:Experimental source: strain IT 4/25/5
C:Genetics:
A:Introns: 2493/3
A:Note: R29R+var1

Query Match      90.2%; Score 37; DB 2; Length 2706;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 PDTRPAP 7
      |||||
Db      2450 PDTRPVP 2456

RESULT 7
G82129
hypothetical protein VC2005 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82129
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82129
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-118 <HEI>
A:Cross-references: UNIPROT:Q9KQJ3; GB:AE004275; GB:AE003852; NID:G9656544; PIDN:AAF9515
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2005
A:Map position: 1

Query Match      85.4%; Score 35; DB 2; Length 118;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 PDTRPAP 7
      |||||
Db      26 PDTPAP 32

RESULT 8
D83142
hypothetical protein PA4040 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83142
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83142
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-335 <STO>
A:Cross-references: UNIPROT:Q9HWY7; GB:AE004820; GB:AE000223; NID:G9950223; PIDN:AAG0742
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4040

Query Match      85.4%; Score 35; DB 2; Length 335;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 PDTRPAP 7  
 Db 308 PDAPAP 314

RESULT 9  
 A13083  
 monooxygenase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 C:Accession: A13083  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: A13083  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-341 <KUR>  
 A:CROSS-references: UNIPROT:Q8U806; GB:AE008689; PIDN:AA145087.1; PID:g17742754; GSPDB:G  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Act4293  
 A:Map position: linear chromosome  
 C:Superfamily: alkanal monooxygenase [FMN-linked] (bacterial luciferase)

Query Match 85.4%; Score 35; DB 2; Length 341;  
 Best Local Similarity 85.7%; Pred. No. 49;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 Db 270 PDTRPAP 276

RESULT 10  
 H98202  
 hypothetical protein AGR\_L1143 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C:Accession: H98202  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: H98202  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-355 <KUR>  
 A:CROSS-references: UNIPROT:Q8U806; GB:AE007870; PIDN:AAK89146.1; PID:g15158956; GSPDB:G  
 C:Genetics:  
 A:Gene: AGR\_L1143  
 A:Map position: linear chromosome

Query Match 85.4%; Score 35; DB 2; Length 355;  
 Best Local Similarity 85.7%; Pred. No. 51;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 Db 284 PDTRPAP 290

RESULT 11  
 E81838  
 hypothetical protein NMA1474 [imported] - Neisseria meningitidis (strain Z2491 serogroup  
 C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
 C:Accession: E81838  
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
 ; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
 Nature 404, 502-506, 2000  
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
 A:Reference number: A81775; MUID:20222556; PMID:10761919  
 A:Accession: E81838  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-365 <PAR>  
 A:CROSS-references: UNIPROT:Q9JU64; GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB847  
 A:Experimental source: serogroup A, strain Z2491  
 C:Genetics:  
 A:Gene: NMA1474  
 C:Superfamily: Neisseria meningitidis hypothetical protein NMB1269

Query Match 85.4%; Score 35; DB 2; Length 365;  
 Best Local Similarity 85.7%; Pred. No. 52;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 Db 8 PDTRPAP 14

RESULT 12  
 G81102  
 hypothetical protein NMB1269 [imported] - Neisseria meningitidis (strain MC58 serogroup  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
 C:Accession: G81102  
 R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.J.  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: A81000; MUID:20175755; PMID:10710307  
 A:Accession: G81102  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-365 <TET>  
 A:CROSS-references: UNIPROT:Q9JU68; GB:AE002475; GB:AE002098; NID:g7226502; PIDN:AAF416  
 A:Experimental source: serogroup B, strain MC58  
 C:Genetics:  
 A:Gene: NMB1269  
 C:Superfamily: Neisseria meningitidis hypothetical protein NMB1269

Query Match 85.4%; Score 35; DB 2; Length 365;  
 Best Local Similarity 85.7%; Pred. No. 52;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 Db 8 PDTRPAP 14

RESULT 13  
 G84302  
 hypothetical protein Vng1484h [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: G84302  
 R:NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
 A:Title: Genome sequence of Halobacterium species NRC-1  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: G84302  
 A:Status: Preliminary

A:Molecule type: DNA  
A:Residues: 1-422 <STO>  
A:Cross-references: UNIPROT:Q9HPT0; GB:AE004437; NID:g10580983; PIDN:AAG19787.1; GSPDB:G  
C:Genetics:  
A:Gene: VNG1484H

Query Match 85.4%; Score 35; DB 2; Length 422;  
Best Local Similarity 85.7%; Pred. NO. 61;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7  
|||  
Db 176 PDTRPAP 182

RESULT 14  
T05839  
subtilisin-like proteinase homolog F17L22.100 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 09-Apr-1999 #sequence\_revision 09-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T05839  
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba  
submitted to the Protein Sequence Database, February 1999  
A:Reference number: Z15454  
A:Accession: T05839  
A:Molecule type: DNA  
A:Residues: 1-769 <BEV>  
A:Cross-references: UNIPROT:Q9SVT3; EMBL:AL035527; GSPDB:GN00062; ATSP:F17L22.100  
A:Experimental source: cultivar Columbia; BAC clone F17L22  
C:Genetics:  
A:Gene: ATSP:F17L22.100  
A:Map position: 4  
A:Introns: 20/1; 79/1; 141/1; 240/1; 363/3; 458/2; 496/3  
C:Superfamily: subtilisin-like proteinase ag12; subtilisin homology

Query Match 85.4%; Score 35; DB 2; Length 769;  
Best Local Similarity 85.7%; Pred. NO. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7  
|||  
Db 429 PDTRPAP 435

RESULT 15  
A43359  
microtubule-associated protein MAP1A - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: A43359; S22108  
R:Langkopf, A.; Hammarback, J.A.; Muller, R.; Vallee, R.B.; Garner, C.C.  
J. Biol. Chem. 267, 16561-16566, 1992  
A:Title: Microtubule-associated proteins 1A and LC2. Two proteins encoded in one messeng  
A:Reference number: A43359; PMID:92355629; PMID:1379599  
A:Accession: A43359  
A:Molecule type: mRNA  
A:Residues: 1-2774 <LAN>  
A:Cross-references: UNIPROT:P34926; GB:M83196; NID:g205537; PIDN:AAB48069.1; PID:g205538  
A>Note: sequence extracted from NCBI backbone (NCBI:111039, NCBI:111040)  
R:Cravchik, A.  
submitted to the EMBL Data Library, June 1992  
A:Reference number: S22108  
A:Accession: S22108  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 73-364, 'OKN', 374, 'PSPKGL', 381-751, 'RSMMSQNNQRR', 764, 'D', 766, 'LR  
'WLKKNMCFQKQSP', 851, 'V', 853, 'NSL', 855, 'LPHRLRLTN', 865, 'W', 867, 'HSQLPDGGD', 877, 'Q', 879,  
A:Cross-references: EMBL:X66840  
A:Experimental source: strain Sprague Dawley  
C:Superfamily: microtubule-associated protein MAP1B  
C:Keywords: microtubule binding; phosphoprotein

Query Match 85.4%; Score 35; DB 2; Length 2774;

Best Local Similarity 85.7%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7  
|||  
Db 757 PDTRPAP 763

Search completed: December 29, 2004, 13:13:00  
Job time : 26 secs

This Page Blank (uspto)



```
; PRIOR FILING DATE: 1999-01-25
; PRIOR APPLICATION NUMBER: 08/542,003
; PRIOR FILING DATE: 1995-10-12
; PRIOR APPLICATION NUMBER: 08/322,760
; PRIOR FILING DATE: 1994-10-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 13
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Antibody Sm3 Complex with Its Peptide Epitope
US-10-371-067-7

Query Match      100.0%; Score 41; DB 15; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
Db 4 PDTRPAP 10

RESULT 7
US-09-822-698A-40
; Sequence 40, Application US/09822698A
; Patent No. US20020146750A1
; GENERAL INFORMATION:
; APPLICANT: Hoogenboom, Hendricus R.J.M.
; APPLICANT: Henderikx, Maria P.G.
; TITLE OF INVENTION: MUCIN-1 Specific Binding Members and Methods of Use Thereof
; FILE REFERENCE: DYX-015.1 US
; CURRENT APPLICATION NUMBER: US/09/822,698A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 09/538,913
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Microsoft Word
; SEQ ID NO 40
; LENGTH: 15
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: PAP peptide epitope of MUC1 bound by PH1 Fab antibody
US-09-822-698A-40

Query Match      100.0%; Score 41; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
Db 1 PDTRPAP 7

RESULT 8
US-09-984-183-17
; Sequence 17, Application US/09984183
; Patent No. US20020142983A1
; GENERAL INFORMATION:
; APPLICANT: AGRAWAL, BABITA
; APPLICANT: LONGENECKER, MICHAEL B.
; TITLE OF INVENTION: MUC-1 ANTAGONISTS AND METHODS OF TREATING IMMUNE
; FILE REFERENCE: 042881/0130
; CURRENT APPLICATION NUMBER: US/09/984,183
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/457,354
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/111,973
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 30

; PRIOR FILING DATE: 1999-01-25
; PRIOR APPLICATION NUMBER: 08/542,003
; PRIOR FILING DATE: 1995-10-12
; PRIOR APPLICATION NUMBER: 08/322,760
; PRIOR FILING DATE: 1994-10-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 13
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Antibody Sm3 Complex with Its Peptide Epitope
US-10-371-067-7

Query Match      100.0%; Score 41; DB 15; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
Db 4 PDTRPAP 10

RESULT 7
US-09-822-698A-40
; Sequence 40, Application US/09822698A
; Patent No. US20020146750A1
; GENERAL INFORMATION:
; APPLICANT: Hoogenboom, Hendricus R.J.M.
; APPLICANT: Henderikx, Maria P.G.
; TITLE OF INVENTION: MUCIN-1 Specific Binding Members and Methods of Use Thereof
; FILE REFERENCE: DYX-015.1 US
; CURRENT APPLICATION NUMBER: US/09/822,698A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 09/538,913
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Microsoft Word
; SEQ ID NO 40
; LENGTH: 15
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: PAP peptide epitope of MUC1 bound by PH1 Fab antibody
US-09-822-698A-40

Query Match      100.0%; Score 41; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
Db 1 PDTRPAP 7

RESULT 8
US-09-984-183-17
; Sequence 17, Application US/09984183
; Patent No. US20020142983A1
; GENERAL INFORMATION:
; APPLICANT: AGRAWAL, BABITA
; APPLICANT: LONGENECKER, MICHAEL B.
; TITLE OF INVENTION: MUC-1 ANTAGONISTS AND METHODS OF TREATING IMMUNE
; FILE REFERENCE: 042881/0130
; CURRENT APPLICATION NUMBER: US/09/984,183
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/457,354
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/111,973
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 30
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-183-17

Query Match      100.0%; Score 41; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
Db 6 PDTRPAP 12

RESULT 9
US-09-984-333-7
; Sequence 7, Application US/09984333
; Patent No. US20020159969A1
; GENERAL INFORMATION:
; APPLICANT: AGRAWAL, Babita
; APPLICANT: LONGENECKER, Bryan Michael
; APPLICANT: REDDISH, Mark Austin
; TITLE OF INVENTION: SMALL PEPTIDE-BASED THERAPEUTICS FOR REVERSING
; FILE REFERENCE: 042881/0116
; CURRENT APPLICATION NUMBER: US/09/984,333
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 09/182,887
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/064,146
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065,209
; PRIOR FILING DATE: 1997-11-12
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-333-7

Query Match      100.0%; Score 41; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7
Db 6 PDTRPAP 12

RESULT 10
US-09-143-379-1
; Sequence 1, Application US/09143379
; Publication No. US20040077826A1
; GENERAL INFORMATION:
; APPLICANT: KOGANTY, R. Rao
; APPLICANT: QIU, Dongxu
; APPLICANT: GANDHI, Sham
; TITLE OF INVENTION: RANDOMLY GENERATED GLYCOPEPTIDE COMBINATORIAL LIBRARIES
; FILE REFERENCE: 042881/0119
; CURRENT APPLICATION NUMBER: US/09/143,379
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: 60/056,240
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

; OTHER INFORMATION: Description of Artificial Sequence: Peptide  
; OTHER INFORMATION: derived from cancer-associated MUC1  
US-09-143-379-1

Query Match 100.0%; Score 41; DB 11; Length 16;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
Db 6 PDTRPAP 12

RESULT 11  
US-09-815-346-3  
; Sequence 3, Application US/09815346  
; Publication No. US20020018806A1  
; GENERAL INFORMATION:  
; APPLICANT: AGRAWAL, BABITA  
; APPLICANT: KOGANTY, R. RAO  
; APPLICANT: LONGENECKER, MICHAEL B.  
; TITLE OF INVENTION: LIPOPEPTIDE ADJUVANTS  
; FILE REFERENCE: 042881-0176  
; CURRENT APPLICATION NUMBER: US/09/815,346  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: 60/191,736  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-815-346-3

Query Match 100.0%; Score 41; DB 9; Length 18;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
Db 6 PDTRPAP 12

RESULT 12  
US-10-106-876-6  
; Sequence 6, Application US/10106876  
; Publication No. US20030157160A1  
; GENERAL INFORMATION:  
; APPLICANT: BUDZYNSKI, WLADYSLAW A.  
; APPLICANT: KOGANTY, R. RAO  
; APPLICANT: KRANTZ, MARK J.  
; APPLICANT: LONGENECKER, B. MICHAEL  
; TITLE OF INVENTION: VACCINE FOR MODULATING BETWEEN T1 AND T2 IMMUNE  
; FILE REFERENCE: 042881-0176  
; CURRENT APPLICATION NUMBER: US/10/106,876  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: 60/278,698  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-10-106-876-6

Query Match 100.0%; Score 41; DB 14; Length 18;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
Db 6 PDTRPAP 12

RESULT 13  
US-10-106-876-17  
; Sequence 17, Application US/10106876  
; Publication No. US20030157160A1  
; GENERAL INFORMATION:  
; APPLICANT: BUDZYNSKI, WLADYSLAW A.  
; APPLICANT: KOGANTY, R. RAO  
; APPLICANT: KRANTZ, MARK J.  
; APPLICANT: LONGENECKER, B. MICHAEL  
; TITLE OF INVENTION: VACCINE FOR MODULATING BETWEEN T1 AND T2 IMMUNE  
; FILE REFERENCE: 042881-0176  
; CURRENT APPLICATION NUMBER: US/10/106,876  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: 60/278,698  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-10-106-876-17

Query Match 100.0%; Score 41; DB 14; Length 18;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
Db 7 PDTRPAP 13

RESULT 14  
US-09-815-346-2  
; Sequence 2, Application US/09815346  
; Publication No. US20020018806A1  
; GENERAL INFORMATION:  
; APPLICANT: AGRAWAL, BABITA  
; APPLICANT: LONGENECKER, MICHAEL B.  
; APPLICANT: PARKER, JOANNE  
; TITLE OF INVENTION: LIPOPEPTIDE ADJUVANTS  
; FILE REFERENCE: 042881-0160  
; CURRENT APPLICATION NUMBER: US/09/815,346  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: 60/191,736  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-815-346-2

Query Match 100.0%; Score 41; DB 9; Length 19;  
Best Local Similarity 100.0%; Pred. No. 5.9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7  
|||||  
Db 6 PDTRPAP 12

RESULT 15  
US-09-815-346-6  
; Sequence 6, Application US/09815346  
; Publication No. US20020018806A1  
; GENERAL INFORMATION:  
; APPLICANT: AGRAWAL, BABITA  
; APPLICANT: LONGENECKER, MICHAEL B.  
; APPLICANT: PARKER, JOANNE  
; TITLE OF INVENTION: LIPOPEPTIDE ADJUVANTS  
; FILE REFERENCE: 042881/0160  
; CURRENT APPLICATION NUMBER: US/09/815.346  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: 60/191,736  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-815-346-6

Query Match 100.0%; Score 41; DB 9; Length 19;  
Best Local Similarity 100.0%; Pred. No. 5.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PDTRPAP 7  
|||||  
Db 6 PDTRPAP 12

Search completed: December 29, 2004, 13:24:26  
Job time : 422 secs

This Page Blank (uspto)

DR WPI; 2002-083037/11.  
 XX New T cell binding ligand peptide for treating immunological disorders  
 PT such as herpes simplex virus, tuberculosis, cancers, acquired  
 PT immunodeficiency syndrome and allergies.  
 XX  
 XX Disclosure; Page 26; 110pp; English.  
 XX  
 CC The present invention relates to novel T-cell binding ligand (TCBL)  
 CC peptides (e.g. peptide G' (modified human MHC class II beta chain peptide  
 CC G' peptide J' (human beta-2-microglobulin peptide) and HIV-1 peptides) and  
 CC TCBL peptide constructs for treating immunological disorders. The peptide  
 CC constructs are useful for eliciting a cellular immune response in a human  
 CC patient. The method comprises administering the peptide construct to the  
 CC patient preferably in combination with an immune response adjuvant. The  
 CC peptide constructs in the form of conjugated peptides are useful for  
 CC eliciting a cellular immune response in a patient exposed to or at risk  
 CC for exposure to the human immunodeficiency virus (HIV). The TCBL peptides  
 CC are useful for treating a patient suffering from an immunological  
 CC disorder such as herpes simplex virus (HSV) infection, malaria,  
 CC tuberculosis, cancers, acquired immunodeficiency syndrome (AIDS),  
 CC allergies, autoimmune diseases (e.g. arthritis, Graves disease, multiple  
 CC sclerosis (MS), autoimmune myocarditis, diabetes and lupus) by  
 CC administering a peptide construct comprising a TCBL peptide bonded to an  
 CC antigenic peptide associated with the disorder. Unlike prior art peptide  
 CC conjugates, a modified version of peptide G has long range stabilisation  
 CC and also enhances the immune response. AAU82019-AAU82114 represent T-cell  
 CC specific binding ligand peptides, peptide constructs or peptides used in  
 CC their construction  
 XX  
 XX Sequence 8 AA;  
 SQ

Query Match 100.0%; Score 41; DB 5; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PDTRPAP 7  
 Db 2 PDTRPAP 8  
 |||||  
 |||||

RESULT 5  
 AAU820933  
 ID AAU820933 standard; peptide; 8 AA.  
 AC AAU820933;  
 XX  
 XX 01-JUL-2002 (first entry)  
 XX Human tumour-associated MUC1 antigen epitope #2.  
 DE Human tumour-associated MUC1 antigen; epitope; cytostatic; vaccine;  
 XX therapy; glycoprotein; immune system.  
 KW Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 3  
 FT /note= "Linked to single alpha-N-Acetyl-galactosamine  
 FT residue in O-glycosidic linkage"  
 XX  
 XX EP1182210-A1.  
 XX  
 XX 27-FEB-2002.  
 XX  
 XX 17-AUG-2000; 2000EP-00117735.  
 XX  
 XX 17-AUG-2000; 2000EP-00117735.  
 XX (BAST/) BASTERT G.  
 XX  
 XX Bastert G, Kaul S;  
 XX

DR WPI; 2002-282784/33.  
 XX Carbohydrate-containing epitopes of tumor-associated MUC1 antigens  
 PT specifically bound by monoclonal antibody 7F11 and 1B4, useful as vaccine  
 PT for treating humans e.g. tumor patient or animal.  
 XX  
 XX Claim 3; Fig 4b; 38pp; English.  
 XX  
 CC The invention relates to epitopes of tumour-associated mucin MUC1 antigen  
 CC comprising an amino acid sequence with atleast one carbohydrate moiety at  
 CC position 3. MUC1 is a transmembrane glycoprotein which is located on the  
 CC luminal surface of ductal epithelial. A tumour vaccine and/or a  
 CC pharmaceutical composition is useful for immunisation in humans (tumour  
 CC patients) or animals and for the production of monoclonal antibodies. A  
 CC diagnostic composition is useful in an diagnostic method useful for  
 CC detection of tumour associated MUC1 antigen within body of a patient. A  
 CC glycopeptide antigen is useful in a diagnostic method for the detection  
 CC of antibodies and/or cells of the cellular or humoral immune system of an  
 CC individual. The present sequence is human tumour-associated MUC1 antigen  
 CC epitope  
 XX  
 XX Sequence 8 AA;  
 SQ

Query Match 100.0%; Score 41; DB 5; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PDTRPAP 7  
 Db 1 PDTRPAP 7  
 |||||  
 |||||

RESULT 6  
 ADL15165  
 ID ADL15165 standard; peptide; 8 AA.  
 XX  
 XX ADL15165;  
 XX  
 XX 17-JUN-2004 (first entry)  
 XX Human tumour-associated antigen, MUC-1 peptide #2.  
 DE Therapeutic binding agent; tumour-associated antigen; MUC-1; cancer;  
 XX tumour; breast carcinoma; colon carcinoma;  
 KW oesophageal squamous cell carcinoma; pancreatic carcinoma;  
 KW prostate carcinoma; multiple myeloma; cytostatic; human.  
 XX  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Region 3..8  
 FT /note= "Specifically claimed in claim 1, and also given  
 FT as SEQ ID No:5"  
 XX  
 XX US6716966-B1.  
 XX  
 XX 06-APR-2004.  
 XX  
 XX 18-AUG-2000; 2000US-00641833.  
 XX  
 XX 18-AUG-1999; 99US-0149492P.  
 PR 11-NOV-1999; 99US-0164714P.  
 XX  
 XX (ALTA-) ALTAREX CORP.  
 XX  
 XX Madiyalakan R;  
 XX  
 XX WPI; 2004-303095/28.  
 XX New binding agent, Alt-1, that binds immunological determinants of MUC-1,  
 PT useful for therapeutically treating a mammal bearing a tumor e.g. breast,  
 PT colon, esophageal, prostate or pancreatic carcinoma, or multiple myeloma.  
 PT  
 XX

PS Claim 1; SEQ ID NO 2; 18pp; English.  
 CC The present invention relates to therapeutic binding agents that bind to  
 CC tumour-associated antigen, MUC-1. The binding agent reduces, reverses, or  
 CC prevents the effects of MUC-1 in cancer. The binding agent is useful for  
 CC treating a mammal bearing a tumour e.g. breast carcinoma, colon  
 CC carcinoma, oesophageal squamous cell carcinoma, pancreatic carcinoma,  
 CC prostate carcinoma or multiple myeloma. The present sequence represents a  
 CC human MUC-1 peptide.  
 XX  
 SQ Sequence 8 AA;

Query Match 100.0%; Score 41; DB 8; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 Db 2 PDTRPAP 8

RESULT 7  
 AAR68021  
 ID AAR68021 standard; peptide; 9 AA.

XX AC AAR68021;  
 XX DT 25-MAR-2003 (revised)  
 DT 05-SEP-1995 (first entry)  
 XX Mucin peptide p20-28.

XX Mucin; multiple tandem repeat; vaccine; pancreas cancer; breast cancer;  
 KW ovary cancer; colon cancer; HIV.

XX Synthetic.

XX PN WO9503825-A1.  
 XX PD 09-FEB-1995.  
 XX PF 29-JUL-1994; 94WO-US008477.  
 XX PR 30-JUL-1993; 93US-00099354.

XX PA (FINN/) FINN O J.  
 PA (FONT/) FONTENOT J D.  
 PA (MONT/) MONTELARO R C.

XX PI Finn OJ, Fontenot JD, Montelaro RC;  
 XX WPI; 1995-082033/11.

XX Synthetic multiple tandem repeat mucin-1 peptides and analogues - have  
 PT native conformation in the absence of glycosylation and are linked to  
 PT epitopes; for vaccines and tests of cancer, viruses and bacteria.

PS Disclosure; Page 54; 125pp; English.

XX A synthetic peptide such as p105 (AAR68022) includes 5 tandem repeats of  
 CC the mucin peptide given in AAR68004. The DTR motif, located between the  
 CC first 2 prolines of each repeat, is the target of an anti-mucin immune  
 CC response, and can be substituted by a sequence from a virus, tumor  
 CC antigen or autoantigen. Drugs based on 9-amino acid portions (AAR68007-  
 CC 21) of the mucin peptide were developed. (Updated on 25-MAR-2003 to  
 CC correct PN field.)  
 XX

SQ Sequence 9 AA;  
 Query Match 100.0%; Score 41; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 Db 2 PDTRPAP 8

RESULT 8  
 AAR68007  
 ID AAR68007 standard; peptide; 9 AA.

XX AC AAR68007;

XX DT 25-MAR-2003 (revised)  
 DT 05-SEP-1995 (first entry)

XX DE Mucin peptide p1-9.

XX Mucin; multiple tandem repeat; vaccine; pancreas cancer; breast cancer;  
 KW ovary cancer; colon cancer; HIV.

XX OS Synthetic.

XX PN WO9503825-A1.

XX PD 09-FEB-1995.

XX PF 29-JUL-1994; 94WO-US008477.

XX PR 30-JUL-1993; 93US-00099354.

XX PA (FINN/) FINN O J.

PA (FONT/) FONTENOT J D.

PA (MONT/) MONTELARO R C.

XX PI Finn OJ, Fontenot JD, Montelaro RC;

XX WPI; 1995-082033/11.

XX Synthetic multiple tandem repeat mucin-1 peptides and analogues - have  
 PT native conformation in the absence of glycosylation and are linked to  
 PT epitopes; for vaccines and tests of cancer, viruses and bacteria.

PS Disclosure; Page 54; 125pp; English.

XX A synthetic peptide such as p105 (AAR68022) includes 5 tandem repeats of  
 CC the mucin peptide given in AAR68004. The DTR motif, located between the  
 CC first 2 prolines of each repeat, is the target of an anti-mucin immune  
 CC response, and can be substituted by a sequence from a virus, tumor  
 CC antigen or autoantigen. Drugs based on 9-amino acid portions (AAR68007-  
 CC 21) of the mucin peptide were developed. (Updated on 25-MAR-2003 to  
 CC correct PN field.)  
 XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 41; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 Db 2 PDTRPAP 8

RESULT 9  
 AAR68007  
 ID AAR68007 standard; peptide; 9 AA.

XX AC AAR68007;

XX DT 11-JAN-1999 (first entry)

XX DE Mucin peptide preparation p 20-28.

XX Mucin; muc-1; human; cancer; infectious disease; vaccine; diagnosis;

KW	multiple tandem repeat; pancreatic cancer; breast cancer; colon cancer.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
XX	US5827666-A.
XX	
PD	27-OCT-1998.
XX	
XX	10-AUG-1994; 94US-00288059.
PF	
XX	
XX	30-JUL-1993; 93US-00099354.
PR	
XX	(UYPI-) UNIV PITTSBURGH.
PA	
XX	Montelaro RC, Fontenot JD, Finn OJ;
PI	WPI; 1998-593988/50.
DR	
XX	Assay for cancer antibodies - using synthetic peptide comprising multiple
PT	tandem repeats of muc-1.
PT	
XX	
PS	Disclosure; Col 25; 45pp; English.
XX	
CC	An assay has been developed for antibodies to pancreatic, breast or colon
CC	cancer in a sample. The assay comprises contacting the sample with a
CC	synthetic muc-1 peptide that comprises at least two 20 amino acid tandem
CC	repeats of muc-1 and is capable of attaining native conformation in the
CC	absence of glycosylation, and detecting any peptide-antibody complex
CC	formation. The assay can be used in the diagnosis of e.g. pancreatic,
CC	breast or colon cancer. The present sequence represents a mucin peptide
CC	preparation from the present invention
XX	
XX	
XX	Sequence 9 AA;
XX	
Query Match	100.0%; Score 41; DB 2; Length 9;
Best Local Similarity	100.0%; Pred. No. 1.7e+06;
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 PDTRPAP 7
DB	2 PDTRPAP 8
DE	AAW72709 standard; peptide; 9 AA.
XX	
XX	AAW72709;
XX	
XX	11-JAN-1999 (first entry)
XX	
DE	Mucin peptide preparation p 1-9.
XX	
XX	Mucin; muc-1; human; cancer; infectious disease; vaccine; diagnosis;
KW	multiple tandem repeat; pancreatic cancer; breast cancer; colon cancer.
KW	
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
XX	US5827666-A.
XX	
PD	27-OCT-1998.
XX	
XX	10-AUG-1994; 94US-00288059.
PF	
XX	
XX	30-JUL-1993; 93US-00099354.
PR	
XX	(UYPI-) UNIV PITTSBURGH.
PA	
XX	Montelaro RC, Fontenot JD, Finn OJ;
PI	WPI; 1998-593988/50.
DR	

```

XX Assay for cancer antibodies - using synthetic peptide comprising multiple
PT tandem repeats of muc-1.
XX
XX Disclosure; Col 25; 45pp; English.
XX
XX An assay has been developed for antibodies to pancreatic, breast or colon
CC cancer in a sample. The assay comprises contacting the sample with a
CC synthetic muc-1 peptide that comprises at least two 20 amino acid tandem
CC repeats of muc-1 and is capable of attaining native conformation in the
CC absence of glycosylation, and detecting any peptide-antibody complex
CC formation. The assay can be used in the diagnosis of e.g. pancreatic,
CC breast or colon cancer. The present sequence represents a mucin peptide
CC preparation from the present invention
XX
XX Sequence 9 AA;
XX
XX Query Match 100.0%; Score 41; DB 2; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 1.7e+06;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX QY 1 PDRPAP 7
XX |||||
XX Db 1 PDRPAP 7
XX
XX RESULT 11
XX AAY05217
XX ID AAY05217 standard; peptide; 9 AA.
XX
XX AC AAY05217;
XX
XX 17-JUN-1999 (first entry)
XX
XX MUC VNTR peptide SEQ ID NO 8.
XX
XX VNTR peptide; mannose receptor; immunoregulatory composition; mucin;
XX immune response; tumour; adenocarcinoma; breast cancer; therapy.
XX
XX Homo sapiens.
XX
XX WO9916455-A1.
XX
XX 08-APR-1999.
XX
XX 29-SEP-1998; 98WO-IB001718.
XX
XX 29-SEP-1997; 97US-0060594P.
XX
XX (AUST-) AUSTIN RES INST.
XX
XX McKenzie IFC, Apostolopoulos V, Pietersz GA;
XX
XX WPI, 1999-254927/21.
XX
XX Immunoregulatory composition comprising mannose receptor-bearing cells,
XX an antigen and mannose, useful for prevention/treatment of cancer.
XX
XX Disclosure; Page 18; 84pp; English.
XX
XX This sequence represents a VNTR from a form of human mucin. The invention
XX relates to an immunoregulatory composition (I) comprising isolated
XX mannose receptor-bearing cells and a conjugate comprising an antigen and
XX a mannose selected from fully oxidised and partially reduced mannose
XX having aldehydes. (I) is useful as a therapeutic agent in animals with
XX natural antibodies against mucin, for inducing a cell mediated immune
XX response specifically to mucin. It is useful for prevention/treatment of
XX tumours, particularly adenocarcinoma, more particularly breast cancer.
XX Prior art methods for removal of tumours uses disfiguring and costly
XX surgery, and/or chemotherapeutic and radiation methods with severe side-
XX effects. The new composition is non-toxic
XX
XX Sequence 9 AA;
XX

```

Sequence 9 AA;

Query Match 100.0%; Score 41; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 DB 2 PDTRPAP 8

RESULT 12  
 AAW67597  
 ID AAW67597 standard; peptide; 9 AA.  
 AC AAW67597;  
 XX  
 XX  
 DT 02-MAR-1999 (first entry)  
 DE T-cell activation peptide #15.  
 KW Activated T helper cell; CD4+; cytotoxic T cell; CD8+; liposome; epitope;  
 KW peripheral blood lymphocyte; antigen-presenting cell; APC; virus; tumour;  
 KW bacterium; parasite; cytokine; vaccine; cancer; malaria; HIV; hepatitis;  
 KW tuberculosis.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9850527-A1.  
 XX  
 PD 12-NOV-1998.  
 XX  
 PF 07-MAY-1998; 98WO-US009288.  
 XX  
 PR 08-MAY-1997; 97US-0045949P.  
 XX  
 PA (BIOM-) BIOMIRA INC.  
 XX  
 PI Agrawal B, Krantz MJ, Reddish MA, Longenecker BM;  
 XX  
 DR WPI; 1999-034715/03.  
 XX  
 PT Method of activation of T cells - by exposure to antigen-presenting cells  
 PT loaded with antigen in liposome, used for, e.g. treating cancer and  
 PT microbial infections.  
 XX  
 PS Claim 14; Page 49; 75pp; English.  
 XX  
 CC Peptides AAW67583-W67611 are used to produce activated T helper (CD4+)  
 CC and cytotoxic (CD8+) T-cells. The activated T cells are produced by  
 CC treating peripheral blood lymphocytes with liposome-encapsulated peptide  
 CC antigen to generate Ag-loaded antigen-presenting cells (APC), contacting  
 CC naive or anergic T-cells with these APC, and isolating the resulting  
 CC activated T-cells. The cells are specific for a particular antigen,  
 CC particularly one derived from a tumour, but also those from viruses,  
 CC bacteria and other parasites. It can also be used to identify antigens  
 CC and epitopes able to generate an Ag-specific T-cell response (by  
 CC assessing proliferation and cytokine release). Also the Ag-loaded APC can  
 CC be used as cellular vaccines for treating cancer (claimed) or other  
 CC diseases (e.g. malaria, human immune deficiency virus infection,  
 CC hepatitis, tuberculosis). The activated T-cells can be used to treat the  
 CC same conditions by adoptive T-cell transfer therapy  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 41; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 DB 1 PDTRPAP 7

RESULT 13  
 AAW67584  
 ID AAW67584 standard; peptide; 9 AA.  
 XX  
 AC AAW67584;  
 XX  
 DT 02-MAR-1999 (first entry)  
 XX  
 DE T-cell activation peptide #2.  
 KW Activated T helper cell; CD4+; cytotoxic T cell; CD8+; liposome; epitope;  
 KW peripheral blood lymphocyte; antigen-presenting cell; APC; virus; tumour;  
 KW bacterium; parasite; cytokine; vaccine; cancer; malaria; HIV; hepatitis;  
 KW tuberculosis.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9850527-A1.  
 XX  
 PD 12-NOV-1998.  
 XX  
 PF 07-MAY-1998; 98WO-US009288.  
 XX  
 PR 08-MAY-1997; 97US-0045949P.  
 XX  
 PA (BIOM-) BIOMIRA INC.  
 XX  
 PI Agrawal B, Krantz MJ, Reddish MA, Longenecker BM;  
 XX  
 DR WPI; 1999-034715/03.  
 XX  
 PT Method of activation of T cells - by exposure to antigen-presenting cells  
 PT loaded with antigen in liposome, used for, e.g. treating cancer and  
 PT microbial infections.  
 XX  
 PS Claim 14; Page 49; 75pp; English.  
 XX  
 CC Peptides AAW67583-W67611 are used to produce activated T helper (CD4+)  
 CC and cytotoxic (CD8+) T-cells. The activated T cells are produced by  
 CC treating peripheral blood lymphocytes with liposome-encapsulated peptide  
 CC antigen to generate Ag-loaded antigen-presenting cells (APC), contacting  
 CC naive or anergic T-cells with these APC, and isolating the resulting  
 CC activated T-cells. The cells are specific for a particular antigen,  
 CC particularly one derived from a tumour, but also those from viruses,  
 CC bacteria and other parasites. It can also be used to identify antigens  
 CC and epitopes able to generate an Ag-specific T-cell response (by  
 CC assessing proliferation and cytokine release). Also the Ag-loaded APC can  
 CC be used as cellular vaccines for treating cancer (claimed) or other  
 CC diseases (e.g. malaria, human immune deficiency virus infection,  
 CC hepatitis, tuberculosis). The activated T-cells can be used to treat the  
 CC same conditions by adoptive T-cell transfer therapy  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 41; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PDTRPAP 7  
 DB 3 PDTRPAP 9

RESULT 14  
 AAW67596  
 ID AAW67596 standard; peptide; 9 AA.  
 XX  
 AC AAW67596;  
 XX  
 DT 02-MAR-1999 (first entry)  
 XX  
 DE T-cell activation peptide #14.  
 XX





This Page Blank (uspto)